

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents *will not* correct images,

please do not report the images to the
Image Problem Mailbox.

THIS PAGE BLANK (USPTO)

SEARCH REQUEST FORM

Scientific and Technical Information Center

Access DB#

54203

Requester's Full Name: _____ Examiner #: _____ Date: _____
 An Unit: _____ Phone Number 30 _____ Serial Number: _____
 Mail Box and Bldg/Room Location: _____ Results Format Preferred (circle): PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need.

 Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc. if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: _____

Inventors (please provide full names): _____

Earliest Priority Filing Date: _____

**For Sequence Searches Only* Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.*

BEST AVAILABLE COPY

STAFF USE ONLY

	Type of Search	Vendors and cost where applicable
Searcher <u>P. Schreiber</u>	NA Sequence (#) <u>1</u>	STN _____
Searcher Phone # <u>308-4272</u>	AA Sequence (#) _____	Dialog _____
Searcher Location <u>CM 112E18</u>	Structure (#) _____	Questel Orbit _____
Date Searcher Provided <u>11/12</u>	Bibliographic _____	Dr. Lina _____
Date Completed <u>11/15</u>	Litigation _____	Lexis Nexis _____
Searcher Prep & Review Time <u>5</u>	Fulltext _____	Sequence Systems <u>CompuGel</u>
Client Prep Time _____	Patent Family _____	WPAWA Internet _____
Indexing Time <u>10</u>	Other _____	Other Specified _____

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 3, 2001, 15:47:29 ; Search time 10414.5 Seconds
(without alignments)
12052.456 Million cell updates/sec

Title: US-09-778-516-1
Perfect score: 8115
Sequence: 1 gatgtacggcgccagatatcac.....gcgtgtcgtctagcagcgttc 8115

Scoring table: IDENTITY_NDC
Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 773874588 residues

Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl: *
1: gb_ba1: *
2: gb_ba2: *
3: gb_ba3: *
4: gb_in1: *
5: gb_in2: *
6: gb_in3: *
7: gb_om: *
8: gb_ov: *
9: gb_pat1: *
10: gb_pat2: *
11: gb_ph: *
12: gb_pl1: *
13: gb_pl2: *
14: gb_pl3: *
15: gb_pl4: *
16: em_ba1: *
17: em_ba2: *
18: em_fun: *
19: em_htgo_hum: *
20: em_htgo_inv: *
21: em_htgo_rod: *
22: em_htg_hum1: *
23: em_htg_hum2: *
24: em_htg_hum3: *
25: em_htg_hum4: *
26: em_htg_hum5: *
27: em_htg_hum6: *
28: em_htg_hum7: *
29: em_htg_hum8: *
30: em_htg_inv1: *
31: em_htg_inv2: *
32: em_htg_other: *
33: em_htg_rod: *
34: em_hum1: *
35: em_hum2: *
36: em_hum3: *
37: em_hum4: *
38: em_hum5: *
39: em_hum6: *
40: em_hum7: *
41: em_in: *
42: em_om: *
43: em_or: *

44: em_ov: *
45: em_pat: *
46: em_ph: *
47: em_pl: *
48: em_rod: *
49: em_sts: *
50: em_sy: *
51: em_un: *
52: em_vi: *
53: gb_sts1: *
54: gb_sts2: *
55: gb_sts3: *
56: gb_sy: *
57: gb_un: *
58: gb_v11: *
59: gb_v12: *
60: gb_htg1: *
61: gb_htg2: *
62: gb_htg3: *
63: gb_htg4: *
64: gb_htg5: *
65: gb_htg6: *
66: gb_htg7: *
67: gb_htg8: *
68: gb_htg9: *
69: gb_htg10: *
70: gb_htg11: *
71: gb_htg12: *
72: gb_htg13: *
73: gb_htg14: *
74: gb_htg15: *
75: gb_htg16: *
76: gb_htg17: *
77: gb_htg18: *
78: gb_htg19: *
79: gb_htg20: *
80: gb_htg21: *
81: gb_htg22: *
82: gb_htg23: *
83: gb_htg24: *
84: gb_htg25: *
85: gb_pr1: *
86: gb_pr2: *
87: gb_pr3: *
88: gb_pr4: *
89: gb_pr5: *
90: gb_pr6: *
91: gb_pr7: *
92: gb_pr8: *
93: gb_pr9: *
94: gb_rod1: *
95: gb_rod2: *
96: gb_in4: *
97: gb_pr10: *
98: em_ba3: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	3535	43.6	3660	3	LBAGAL	M23530 L.delbrueck
2	3497.2	43.1	5059	9	AX009488	AX009488 Sequence
3	3386.6	41.7	5015	3	LBALAC2BUL	M55068 L.delbrueck
4	1336.4	16.5	3840	56	EVEI32038	AJ132038 Expressio
5	1240.4	15.3	3853	9	AR098190	AR098190 Sequence
6	1238.4	15.3	2093	3	LP1PLPREP	M31223 Plasmid pLP
7	1193.8	14.7	2140	2	C300RFR	J03319 Plasmid pC3
8	1082	13.3	3986	56	PCDNA32EO	X90639 Cloning vec

9	1068.6	13.2	4026	9	AR098191	AR098191	Sequence
10	988	12.2	4597	10	AX060344	AX060344	Sequence
11	973	12.0	4283	10	163120	163120	Sequence 3
12	973	12.0	4283	10	185496	185496	Sequence 1
13	968.6	11.9	4050	56	AF053407	AF053407	Expressio
14	968.6	11.9	4227	56	AF053408	AF053408	Expressio
15	968.6	11.9	4625	56	AF053409	AF053409	Expressio
16	968.2	11.9	4657	56	AF053406	AF053406	Expressio
17	959	11.8	5732	56	AX027904	AX027904	Sequence
18	959	11.8	5732	56	AX027910	AX027910	Sequence
19	935.6	11.5	5432	9	AX026821	AX026821	Sequence
20	935.2	11.5	5729	9	AX027908	AX027908	Sequence
21	935.2	11.5	5729	56	AX027914	AX027914	Sequence
22	928.6	11.4	6094	9	AX027906	AX027906	Sequence
23	928.6	11.4	6094	56	AX027912	AX027912	Sequence
24	921.8	11.4	6285	9	AR018923	AR018923	Sequence
25	921.8	11.4	6285	9	AR066551	AR066551	Sequence
26	921.8	11.4	6285	10	172268	172268	Sequence 49
27	921.8	11.4	6285	10	176835	176835	Sequence 49
28	909.2	11.2	6525	56	AF063848	AF063848	Cloning v
29	909.2	11.2	6525	56	AF063585	AF063585	Cloning v
30	909.2	11.2	6706	56	AF063850	AF063850	Cloning v
31	909.2	11.2	7758	56	AF063849	AF063849	Cloning v
32	908.2	11.2	4109	56	PSPORT1	PSPORT1	Cloning v
33	907	11.2	3018	56	AF324725	AF324725	Cloning v
34	907	11.2	3141	56	AF324726	AF324726	Cloning v
35	906.8	11.2	7511	56	CVU37023	CVU37023	Cloning vec
36	906.8	11.2	7511	56	CVU37068	CVU37068	Cloning vec
37	906.4	11.2	2000	10	103343	103343	Sequence 3
38	906.4	11.2	2661	56	SYN8MRG	SYN8MRG	Sequence 3
39	906.4	11.2	2689	56	AF016541	AF016541	Synthetic P
40	906.4	11.2	2704	9	AX018504	AX018504	Cloning v
41	906.4	11.2	2820	56	AF222336	AF222336	Sequence
42	906.4	11.2	2860	56	AF017063	AF017063	Cloning v
43	906.4	11.2	2999	56	SYNCCDBA	SYNCCDBA	Cloning vec
44	906.4	11.2	2999	56	SYNCCDBB	SYNCCDBB	Cloning vec
45	906.4	11.2	3024	3	PK9CDB	PK9CDB	L27082 Plasmid KIL

ALIGNMENTS

RESULT	1	LBAGAL	3660 bp	DNA	BCT	26-APR-1993
LOCUS		L.delbrueckii bulgaricus beta-galactosidase gene, complete cds.				
DEFINITION		M23530				
ACCESSION		M23530.1	GI:149546			
VERSION						
KEYWORDS		galactosidase.				
SOURCE		L.delbrueckii bulgaricus (strain B131) DNA.				
ORGANISM		Lactobacillus delbrueckii				
		Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillaceae;				
		Lactobacillus.				
REFERENCE		1 (bases 1 to 3660)				
AUTHORS		Schmidt, B.F., Adams, R.M., Reguadt, C., Power, S. and Mainzer, S.E.				
TITLE		Expression and nucleotide sequence of the Lactobacillus bulgaricus				
JOURNAL		beta-galactosidase gene cloned in Escherichia coli				
MELINE		J. Bacteriol. 171, 625-635 (1989)				
COMMENT		Draft entry and computer-readable sequence for [1] kindly submitted				
		by B.F.Schmidt, 29-MAR-1989.				
FEATURES		Location/Qualifiers				
SOURCE		1. 3660				
		/organism="Lactobacillus delbrueckii"				
		/db_xref="taxon:1584"				
		452. 458				
		/note="ribosomal binding site"				
		468. 3491				
		/note="beta-galactosidase"				
		/codon_start=1				
		/transl_table=1				
		/protein_id="AAA25240.1"				
		/db_xref="GI:149547"				

RBS

CDS

/translation="MSNKLKREKVDQADLAMLTDPEVEYNTPHSDHESFOSQER	
LEEGKSSIVSDLDWLDIDYENEGQVNFPAEDSDSNFESVAPGLQLGQGPQ	
LYVYQPMDSSEIEFPDIPSKNPLASVYFDDLEFMDKEVLKFDGATATAYWL	
NGHFVGESEFSPSEFMVTRFKLKENNRILAVALKSSASMLDODPMRSGLEFV	
TLQAKPHLHEDLTLASLDNKKOKGKLVNAYLAPNSFLKLEVRSDGLVAEK	
LGPTRSEDLFTLADLPEYMAASAKRPNIYQRLYLQAGLSLEVRQGRNELD	
GIWYLNQRIYFKANRHEFDSKLGRLATEEDMDLWIKTKMRSLINVRCSHYNQSL	
FYRLCDXGRLVDEIDANLESHGTGKVGSHRDPNVDQHWLGASLSVKMMAR	
DKNHSILIMELGNEASYAGVFAQMDVFRKADPTRVQHYEGVHNKRFQDADQIESR	
MYAPAKVLEETLNKPAKPFISVYAHMGVSVDLAATLEYKPHYOGSFLIMID	
OGLEKDHLLYGGDDRPDPEFCGNGVADRDRESKLANVALYANLKLKYDDGO	
LFLKNDNLFTNSSSYFLPILSLVDGKLLTYOSRPLTTCLEPESSTPALPHEVADK	
EVVYRVTAHLKEDLPMADEGTVAAEAEVAKLPDEFKPGRPDLVSDNLGLGNNE	
QILFSEKVGWVSLKTAGREYLRKLPETFEWRLDNDRGAGYDILARWENAKYAR	
LKDISCEYKEDSVLVKTAFTLPVAKLGDLYTVYVDGSKIAVADPEGAEGALLPA	
FGMLALPKELTDYRYGIGLGNESYPTDSENGYIGVAGAKNPSYPRPDTGRSK	
VRYQULDERKGLFETANGALNLNLSALPYSNAQLEADHAEELNNTWVALSAQK	
VGDDSWGKVPHEFCIDQAKARQRLRLOPLFLK"	
3477. .3531	
misc_feature	
/note="region of dyad symmetry"	
BASE COUNT	941 a 927 c 996 g 796 t
ORIGIN	1bp upstream of HindIII site.
Query Match	
Best Local Similarity 98.8%; Pred. No. 0;	
Matches 3590; Conservative 0; Mismatches 0; Indels 45; Gaps 1;	

QY	4480	ttggcagcagctctccgagcaatgaacttgacatggttgagaccggtcttcgc	4539	
DB	11	TTGGCAGCCAGCTCCCGGCAATTATGACATGTGACAGACCCGCTCTTG-69		
QY	4540	aagccgaattgcaccacactgcgcgcgtactagtgatcatcgatgaataaagttaag	4599	
DB	70	-----CCGATAAAAAGTTAG-85		
QY	4600	cgacgagcttgcctcctggtgcccagcagcgttaaggtctacgcgcatttgcgcgtact	4659	
DB	86	CGACGGCTTGGCCCTGCTGCTCCAGACAGCGTAAAGGTCTACGGCCATTCTGT	145	
QY	4660	ccgcagcctgcccagaagaccgagcactcgatcgctctggaataatggatgggtctt	4719	
DB	146	CCGCCAGCTGGCCCAAGACCCGCGACTCGATCGCTCGGAAATGAAACATGGGCTTGGT	205	
QY	4720	cttgattcaacttgccttgcctgagcagcgttaaatgaacggagcttgccttgcgtact	4779	
DB	206	CTTGATTCACTTGGCTTGGCTGGCACGGTCAATTAACGGGACTGGCTTGTCACTATGT	265	
QY	4780	tgaagagggcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc	4839	
DB	266	TGAAGGGCCACAGGTGTAAGGCCGGCCAGCAGATCTCGAATTTCTGGACCCGGCAT	325	
QY	4840	caagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc	4899	
DB	326	CAGACAGGGCCAGCTGAGACAGCAGCGTATCTGACCGCTCAACAGGAACTTTTCAC	385	
QY	4900	aaatgacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc	4959	
DB	386	AAATAGCCAGATGCTTTCGCGATCGGCGACAGCTCCAGCCCTGAGATGATATCA	445	
QY	4960	gttaagaggaagaatagaataatgagcaatagtagtaataaagaataaagaattgacca	5019	
DB	446	GTTAGAGGAGGAATTAATAAATGAGCAATAGTATGATTAATAAAGAAAGATTGACCA	505	
QY	5020	ggcagacgctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc	5079	
DB	506	GGCAGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	565	
QY	5080	cttcgacacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc	5139	
DB	566	CTCCGACCATGAGTCTTCTCAAGCCAGGAGCAACTGGGAGGAGGCAATCTCCAGTTAAT	625	
QY	5140	gagtcctccgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc	5199	

|||||
Db 626 GCAGCCCTGGACGGGGGCGACTGGCTGATGACTACCTGAAAAACGGCGACGAGCAGCTCA 685
OY 5200 ctctatctcagaagaagcttggacgatagaatcttaagtcacatgaagaccggaacct 5259
Db 686 CTCTATCTCAGAAAGACTTTGACGATGACAAATTTTAAGTCAGCAAAAGTAACCGGCAACCT 745
OY 5260 ggaactgcaagagcttggccagcccaatctatctcaagctccaataatcatatggagcgag 5319
Db 746 GGAAGTGAAGGCTTTGGCCAGCCCACTATGTCACAGTCCAAATATCCATGGGAGCGAG 805
OY 5320 tgaagagaattcccgcccaaatcccaagaanaaalcgcgtcgtcttctatgttcagata 5379
Db 806 TGAGAGAGATTTCCCGCCCCCAATTCGAAGCAAAAATCCGCTCGCTTCTTATGTCAGATA 865
OY 5380 cttgacctgtagaagcttctgggagaaagaagtcagcttgaaagtttcagcgagcg 5439
Db 866 CTTTGACCTGGATGAAGCTTTCTGGGACAAAGAACTGAGTTGAAGTTTGAAGCGGGCGCG 925
OY 5440 aacagccatctatgtctgctgaaacgagccacttgcgtgctcagggggaagacctcttac 5499
Db 926 AACAGCCCTATATGTCGTGAGAACGGCCACTTCGTGGCTACGGGGAAGACTCCTTATC 985
OY 5500 cccaagcgagtttatgtgtacaaagttcctcaagaagaanaataaacgccttgcagtgagc 5559
Db 986 CCCAAGCGAGTTTATGTTTACCAAGTTCTCCAGAAAGAAATATACCCGCTGGGAGTGCGC 1045
OY 5560 tctctacaagatctctccgcctcctgcgtgcgtggaagacaggaactctgcgcagtctcg 5619
Db 1046 TCTCTACAAAGTATTCCTCCGCTCTCTGGCTGGAAGACGACACTTCGGCCATGCTCGG 1105
OY 5620 ttgttcaagatcagtgactcttcaggcccaagccgcttgcacttggagagacctaaact 5679
Db 1106 TTTGTTCAGATCAGTGACTCTTTCAGGCCAACCCGGTGTGCACTTGGAGCACTTAACT 1165
OY 5680 tacgagcagcttgcacgaataactcaaaaagaagaagctggaagctcgaagccaatctgc 5739
Db 1166 TACGGCCAGCTTGACCGGTAACCTACCAAAAAGGAAGCTGGAAGTCGAAGCCAAATATTC 1225
OY 5740 ctacgcgtctccaaaatgcacagctttaaagctggaagtcggaagtagtgaagtgactgt 5799
Db 1226 CTACCGGTTGGCAAAATGCACACTTTTAAGCTGGAAGTGGGAGTATGTAAGCTTGTGT 1285
OY 5800 tgcctgaagaagcttggcccaatcagaagcgaagcagctgaatctcctggttgaattggcc 5859
Db 1286 TGCTGAAAAAGCTGGGCCCAATCAGAAAGGACAGCTGGAATTCCTGCTGATTTGCC 1345
OY 5860 agtagctgcctggaagcgcggaanaagccttaaccttaacaggtccgcctgtatattaca 5919
Db 1346 AGTAGCTGCGTGGAGCGCGGAAAAAGCCTTAACCTTACAGGTCCGCGTATTTATACCA 1405
OY 5920 ggcagcgagcctctagaagtttagcggaagtgaggttccgcgaacttgaactaa 5979
Db 1406 GCGCGGAGAGCCTTTAAGAGTTAGCCCGGAGGAAGTGGGTTCCGCAACTTTGAACCTTAA 1465
OY 5980 agacgggattatgaccttaacggccagcgagatgcttcttaaggggcccaacggcgcg 6039
Db 1466 AGAGCGGATTTATGTAACCTTAAAGCGGCGAGGATGCTCTTCAAGGGGGCAACCGGACGA 1535
OY 6040 atttgaagtaagttggtgcgggctatcaacgaagaagagatagtatctgggaacatcaagac 6099
Db 1526 ATTGGAAGATGATTTGGGTGGGCTATCAGGAGAGAGATGTATGTGAGATCAAGAC 1585
OY 6100 catgaagcgaagcaacataatgtctgcgtctgtctcactcaacccgaacagtcctctc 6159
Db 1586 CATGAACGGAAGCAACATCATATGTCGTGCTGCTCTACATCCGAACAGTCCCTCTT 1645
OY 6160 ttacggcctctgtagaagaagtccttcaagctcatgataagcttaacccgaagaagcca 6219
Db 1646 TTACCGGCTCTGTACAAAGTAGCGGCTTTACGTATGATGAACCTTAACCTGGAAAGCCA 1705
OY 6220 cgggacctgggaaaaaagttgggggggacgaagaatcctagctcaatgttccagcgagta 6279
|||||

Db 1706 GGGCAGCTTGGGAAAAAGTGGGGGGGCGACGAAGATCTAGCTTCAATGTTCCAGCGGATGA 1765
OY 6280 ccagatctgcttggagagccagcttaccgggtgagaagaatgataagcttcgggaaagaa 6339
Db 1766 CCAGATTGGCTGGGAGCGACGCTTATCCGGGTGAAGAAACATGATGGCTGGGACAAAGA 1825
OY 6340 ccatgctcaatcccaatctgtctttaggaatgaagcttaacggcgagctgtcttc 6399
Db 1826 CCATCTTCAATCCCAATCTGCTTTTAGGCAATAGTCTTACGCCGCGCAGCTGTTTGGC 1885
OY 6400 ccaaatgtctatcagtcgcggaaagctgataccgacccgggttccagcataagaagggt 6459
Db 1886 CCAATGGCTGATTAATGTCGGAAGCGATGATCCAGCCGGGTTCAAGCTATGAAGGGGT 1945
OY 6460 gaaccaacaacggaaagtttagacgacgccaacagatgtaaagccgagatgtatgtccgc 6519
Db 1946 GAACCAACAACGGAAATTTGAGAGACGCCACCGAGTTAAAGCCGAGTATGCTCCGCG 2005
OY 6520 caaggtatgaagaatacttgacaataaaccgaagccattatctcaagttgaata 6579
Db 2006 CAAAGTTATGAAGATCTTGAACCAATTAACCAAGCCAAAGCCATTATCTCAAGTTGAATA 2065
OY 6580 cgtcacgcacatgycgaactcgtcgtgagcttgcgcgtacacgycgccttggaaaaaata 6639
Db 2066 CGCTACGCCCATGGGCAACTCGCTGGTGAAGCTGCGCCCTACACGGCCCTGGAAAAATA 2125
OY 6640 cccccaactacaaggcggtctcaatctggagctgattgaaccaaggaactgaaaaaagcgg 6699
Db 2126 CCCCACATCAACGAGCGGCTTCAATCTGGAGCTGATTAACCAAGCAAGCAAAAAAGACGG 2185
OY 6700 gcaactgtcttatggggcgagcttcgatgacggcccaacgacatgatatcttcgggaa 6759
Db 2186 GCACCTGCTTATGAGGGGCACTTGATGACCGGCAACCAACCATGTGATTTCTGGGGAA 2245
OY 6760 cgcctgtcttcttgtagccgagctgaatcgcgcgaacttgcgaagtcaatgtcaaggccctta 6819
Db 2246 CGGCTGCTTCTTGCTGACCGGACTGATCGCCGAATCGGCTATGTCAAGGCCCTTTA 2305
OY 6820 cgcaaaccttaagttagaagtaaaagatgggacgctcttcctcaaaaaagcaaatatt 6879
Db 2306 CGCCAACTTAAAGTTAGAGTAAGAAAGATGGGAGCTCTTCTCAAAAACGCAATTTAT 2365
OY 6880 taccacagctcatcttaactctcttctgaagcttcttggatgaggaagttgacctta 6939
Db 2366 TACCAAGCTCATCTTACTTCTTGTGACTGATCTTTTGTGCAATGGCAAGTTGACCTA 2425
OY 6940 ccaagacggcctctgaaccttgccttgcgtgcgtgcgaactcgcggagaccttgcctgccc 6999
Db 2426 CCAGAGCGCGGCTCTGACCTTTGGGCTGAGCCTGGGCAATCCGGACCTTTGGCCCTGCC 2485
OY 7000 ttggccggaagtcgtatgaanaaagggaagtcgctacccgggtlaacgcccactaa 7059
Db 2486 TTGGCCGGAAGTCTGATGAAAAAGGGGAGGTGCTACCGGGTAAACGGCCCACTTAA 2545
OY 7060 aagaagcttgccttggcgagtgagggttactcgttgagcttaacagaagaagtagctca 7119
Db 2546 AGAAGACTTGTGCTTGGGGGAGTGAAGGCTTACGTGCTGAACGAGAAAGTAGCTCA 2605
OY 7120 aaaagctccggaatttaagccggaaggcgcgacagatttgaattgatccgactcaacct 7179
Db 2606 AAAGCTGGCGGAATTTAAAGCGGAAGGGCGCGAGATTATGTAATCCGACTCAACCT 2665
OY 7180 aggccttgaaggaataacttccaaatctcttctccaaggtcaaggcttgcgggttc 7239
Db 2666 AGGCTGAAAAAGAAATTAATTCCAAAATTCCTTCCAAAGGTCAAGGGCTGGCGGTTTC 2725
OY 7240 cctcaagatgctggtaggaagaatacttgaagcgagctccggaatttacccttcgagggc 7299
Db 2726 CCTCAAGTATGCCGGTAGGGAATTAATGGAAGCGGCTGCCGAATTTACCTTCTGGCGGC 2785
OY 7300 cctgaacgacaacgacccgggagctgtttagcgatgatcttgcgagcttgggaaatgac 7359
Db 2786 CTTGACGAGCAACGACCGGGAGCTGGTTAGGCTATGATCTGCGCGGTGGAAAAATGC 2845
|||||

[illegible]

source	1..5059 /organism="Lactobacillus delbrueckii subsp. bulgaricus" /db_xref="taxon:1585"
CDS	122..1873 /note="unnamed protein product" /codon_start=1 /transl_table=1 /protein_id="CAC07488.1" /db_xref="GI:9966772"
CDS	1877..4519 /note="unnamed protein product" /codon_start=1 /transl_table=1 /protein_id="CAE07488.1" /db_xref="GI:9966773"
CDS	/translation="MSNKLVEKRYDDADLAWLTDPEYEVNTTPSHDHSPOSQEE LEEKSLSVOSLDGDMDLIDVAENQSPVNEAEFDSDSESVAPGMLLEVDGGQPQ LYNQVPWDSGEIEFPQIPSKNPASVRVFDEDEFMWDEKFLDKGAATAYTML NGHVGGDESFPTSEFPAVTFFLKNNRLAVALYKXSASMLEDDOFMRSGLFRES TLQAKPRHLEDIKLTSLINDYOKKLEVANIAYRLPNASFLEYRDSGDLVAEKI LGPIRSDELFTLADLPYAAMSARKPNITQYRLITYOAGSLLBESROEPGRNELKD GITMNLQRIYFKGANRHEPDKLGRAITTEDMTDKTKRNSINAVRSCHYRNOST FYRLCDKYGLVIDEAMNLESHGTKEGVGHEDPSFNPGDDHWLGNASLVKNMMAR DKNASILIMSLGENSYAGIVFAQMADYVRKADPTVOHQHYGHNRKFDPDATIIES MYAAKPAIEEYLITNKPAPKAFISVEYAHMGVSVDLAYTAYLEKYPHOGGFIMDIJ QLDSKVGMVPSLKTAYGREYLRKLPETFEWRALTNDNRGAGYGTYDLARMENACKVAR LQIDSCEVKEDSVLKTAFETPLPALKGLDIVTVYEVDGRGKIATVADPFGEAEGLLPRA FGMLALPKELTIDRVRYGGLGPNES"
BASE COUNT	1229 a 1341 c 1323 g 1166 t
ORIGIN	
Query Match	43.1%; Score 3497.2; DB 9; Length 5059;
Best Local Similarity	99.5%; Pred. No. 0;
Matches 3518; Conservative	0; Mismatches 13; Indels 3; Gaps 1;
OY	4584 ccgataaaaagttaagcgacggcgttcctgccttggtgccagcagaacgylaagcttaccgcgc 4643 DB 1479 CCGACAAAAAGTTAGGCAGCGGCTTGCCCTGTGCCACAGCAGTAGGTACTACGGCC 1538
OY	4644 caattgcggaatctccgcacagcttgccacaagaccggacatcgatcgctccggaatatg 4703
DB	1539 CATTTGCCGCTACTCTCCGCGACAGTCGGCCMAAACCCGGGACACGATCGCTCGGAAAATG 1598
OY	4704 aaacatggagcttgtcttgatccaaccttgaggcctgggcaacgyltaaataaacggagact 4763
DB	1599 AACATGGGGCTTGGCTCTTGATTACCTTTGGCCTTG6GCACGGTCAAATTAACGGGACTG 1658
OY	4764 gctttgcacgatlttgaaagagggcacgtaglaaagccggcagcagacagatcctggaat 4823
DB	1659 GCTTTGTACGATATTTGAAGAAGGGCAGCAGAGTAGAAGCCGGCCAGCAGATCTGGAAT 1718
OY	4824 tcctgggaaccggcgatcaagacggccaagctygagcaacaggttaatctgtgacgtcatca 4883
DB	1719 TCTGGGACC CGCGCATCAAGCAGGCCAAGCTGAGACACGGTAAATCTGACACGTCATCA 1778
OY	4884 acagagaaacttcaaaaaatagcagaatgcttccgatactggcagacagcgtccaagccc 4943
DB	1779 ACAGCGAAACTTTCGCAAAATAGCCAGATGCTCTTGGCCGATTCGGCCACACGCTCCAAGCCC 1838

QY 4944 tggatgagtatccaagttagaagggaagaattagaanaattgagcaataagttagtaaaag 5003
|||||
Db 1839 TGGATGATGATTTCAACTTTGAAAGGGAATTTAGAAATGACAAATAGTTACTTAAG 1888
QY 5004 aaaaaaagttagaccagcgagacctggcctgctgactgacccggaaagttacgaagtc 5063
|||||
Db 1899 AAAAAAGTTGACGAGGAGACTTGGCGTGGCTGACTGAGCCGGAGTTTACGAATCA 1958
QY 5064 atacaattcccccgacctccgaccaagagtccttccaagccaggaagaaacttgaaggag 5123
|||||
Db 1959 ATCAATTTCCCGCACTCCGACCATGATGCTTCCAAAGCCAGGAAGAACTGGAGAGG 2018
QY 5124 gcaagtcacatttagtgcagtccttcggacggagacgtgcgattgactgaacgtgaagaag 5183
|||||
Db 2019 GCANGTCATTATGATGCACTCCCTGGACGGGAGCTGGCTGATTGACTGACCTGAAGACG 2078
QY 5184 gccagggaccagctcaactctctatgcaagaagacttgaacgaatgaactttaagtcagtc 5243
|||||
Db 2079 GCCAGGGACGACGTCATCTTATGCAAGACCTTGGACGATGACAAATTTTAAGTCAGTCA 2138
QY 5244 aagtaaccggcaaccttgaacctgaagcttggccagcccgatgtcaacgtccaat 5303
|||||
Db 2139 AAGTACCCGCAACTGGAACTGCAAGGCTTTGGCCAGCCCAAGTATGTCACGTCCAAT 2198
QY 5304 atccatggagcgagtgatgaggaagatcttccgcgcccccaattccaagaaaaatccgtcg 5363
|||||
Db 2199 ATCCATGGGACGAGAGAGAGATTTTCCCGCCCAATTTCCAAAGAAAAATCCGCTCG 2258
QY 5364 ctctctatgcaataacttgaactgtagatgaactctctggaagaagatcagctga 5423
|||||
Db 2259 CTCTTATATGCATATCTTGGACTGATGAGCTTTCTGGACAAAGAGTCAAGCTTGA 2318
QY 5424 agtttgcggggcggaacagcaactatgctgctgagaaogccactgcgtcgtaag 5483
|||||
Db 2319 AGTTTGACGGGGGCGCAACAGCCATCATGTGCGTGAAGCGGCACTGTGCGGTACG 2378
QY 5484 gggaagaactccttaccccaagcgagttaatggttaccagaatctcctaagaagaataa 5543
|||||
Db 2379 GGAAGACCTCTTACCCCAAGCGAGTTTATGTTACCAAGTTCCTCAAAAAATA 2438
QY 5544 aacgcctgagctgagctcctacagaatctctccgcctcctgctggaagaccagact 5603
|||||
Db 2439 ACGGCTGGAGAGTGGCTCTTACAGATCTTCCGCTCTGCTGGAAGACCAAGACT 2498
QY 5604 tctggcagatgctggttctgaagtcagtgactcttcaggccaagccgcgtctgact 5663
|||||
Db 2499 TCTGGGCGCATGTGCTGTTGTTCAGATCAGTCACTTTCAGGCCAAGCCGCTGCAC 2558
QY 5664 tggagaagacttaagcttaagcgagcttgaacgataactccaaaaagaaagcttgaag 5723
|||||
Db 2559 TGGAGGACCTTAAGCTTACGGCAGCTTGAACGATTAACCAAAAAAGAAAGCTGGAAG 2618
QY 5724 tccaagccaatattgctcctacgccttgcacaatgccagcttaagctgaagtgaggata 5783
|||||
Db 2619 TCGAAGGCATATTTGCTTACCGCTTGCCAATGCGACCTTTAAGCTGGAAGTGGGATA 2678
QY 5784 gtagaagtgacttggttctgtaaaagcttggcccaatcagaagcgagcttgaatctca 5843
|||||
Db 2679 GTGAAGGTGACTGTTGCTGTAAGAGCTGGCCCAATCAAGAGGAGACACTGGAATTCA 2738
QY 5844 ctctgctgatttggcagtgactgctcgaggcggaagaaagcttaaccttaccagttcc 5903
|||||
Db 2739 CTCTGGCTGATTTGGCAGTAGTGCCTGGAGCCGGAAAAAGCCTTAACCTTACCAAGTCC 2798
QY 5904 gctctgattatataccagcgagcgagcctcttaaggttgaagctgaagaaagtggttcc 5963
|||||
Db 2799 GCTGTATTTATACGAGCGAGGAGCCTCTTAAAGTTTACCGGAGAAAGTGGGTTCC 2858
QY 5964 gcaactttgaactaaagacgggattatgtacacttaacggcgagctgcgtctcaag 6023
|||||
Db 2859 GCAACTTTGAATAAAGAGAGGATATGTACCTTAAAGCGGAGGATGCTTCAAG 2918

QY 6024 gggccaacccggcacgaatttgcagtaagttggtcgggtatcacggaagaagatga 6083
|||||
Db 2919 GGGCCAAACCGGACAGATTTTACAGTAGTTGGCCGGGCTATATACAGAAAGATATGA 2978
QY 6084 tctggacatcaagaacatgaagcgaaacatcaatgtctgcgtctcaactacc 6143
|||||
Db 2979 TCTGGATATCAAGACCATGAAAGGAAAGCAATCATATGCTGTCCGCTCTCACTACC 3038
QY 6144 cgaacagtcctccttcttaccggtctgtgcaagtaacagccttgaatgaag 6203
|||||
Db 3039 CGAACCAAGTCCCTCTTTTACCGGCTGTGCAAGTACAGGCTTTTACGTCAATGATGAG 3098
QY 6204 cttaaccctggaaagccagcgagccttgggaaaaagtgggggggcgcaagaatctcagttca 6263
|||||
Db 3099 CTAACTTGGAAAGCCACGGCACTTGGGAAAAAGTGGGGGGGACGAAGATCTAGCTTCA 3158
QY 6264 atgttccaagcgatgacccaagcatgtgtggaagcaacttaccgggttgaagaacatga 6323
|||||
Db 3159 ATGTTCCAGGGGATGACAGATTTGGCTGGGAGCCAGCTTATCCGGGTGAAGAAATGA 3218
QY 6324 tggctcgggacaagaacatgcttcaatcctaattcgtctttaggaatgagttcag 6383
|||||
Db 3219 TGGCTCGGACAAAGAACCATGCTTCAATCTGATCTGCTTTTGGCAATGAGTCTTACG 3278
QY 6384 ccggcactgtcttggcccaaatggcttgaagtcgggaaggtcgatccgaccgggttc 6443
|||||
Db 3279 CCGGCACTGTCTTTGGCCAAATGGCTATTAACGTCGGAAGGCTGATCCGACCCGGGCTC 3338
QY 6444 agcacatgaaggggtgacccaacacgaagtttgaacgaacgccaacgaattgaagacc 6503
|||||
Db 3339 AGCAGTATGAAGGGGTGACCCACAAACGGAAGTTTGAAGAGCCACCAAGTGAAGGCC 3398
QY 6504 ggaatgctgcgcggccaaggttaattgaagaatacttgaaccaataaacgccaagccat 6563
|||||
Db 3399 GGATGATGTCTCCGCGCAAGGTATTAAGAAATCTTGACCAATTAACAGCCAAAGCAT 3458
QY 6564 ttactcagtttgaatagctctacggcaatgggcaactcgtcgtgacttggccgctaca 6623
|||||
Db 3459 TTATCTCAGTTGAATGATGCTACGCCATGGCAATCCTCGTGAGTCCGCTTACCA 3518
QY 6624 cggccctggaaaaaatcccccaactacagagcggtctcatcttgaatgagattgaacaa 6683
|||||
Db 3519 CGGCGCTGGAATAATFACCCCACTACAGGGGCGCTTCACTGGAGATGATTAACCAAG 3578
QY 6684 gactggaaaaaagacgggacacttcttaatggggcgacttcgatgacggcgcaacgact 6743
|||||
Db 3579 GACTGGAATAAAGAGCGGCGACTGCTTATGGGGCGACTTGCATGACGGGCAACCCACT 3638
QY 6744 atgaattctgcgggaagcgactggcttcttgaacggagactgaatcgccgaactgcta 6803
|||||
Db 3639 ATGAATTTCTGGGGAACGGCTGGCTTTCCTGTGACCGGACTGAATTCGCCGAATGGCTA 3698
QY 6804 atgtcaagcccttlaagcgcaacacttaagttagaagtaaaagatgggcaactctctcca 6863
|||||
Db 3699 ATGTCAAGGCGCTTACGCCAACCCTTAAGTTAAGATAAAGATGGGACGCTTCCCTCA 3758
QY 6864 aaaaacgaacttaatttaccacaagctcatcttactactcttctgaactagtccttggctg 6923
|||||
Db 3759 AAAACGACAATTTATTTACCAACAGCTCATCTTACTTCTTACATGACTTTTGGCTCG 3818
QY 6924 atggcaagttgacttaccagaagcgagccttgaacttggccttggaaacccctgggaactcg 6983
|||||
Db 3819 ATGGCAAGTTGACCTTACCAAGCGGCGCTGTGACTTGTGGCTGGAACCTGGCAATTCG 3878
QY 6984 ggaacttggcccttgccttggccgggaagtcgtaagtaaaaaaggggaagtcgtctaccggg 7043
|||||
Db 3879 GGAAGCTTTGGCCCTTGGCTTGGCCGGAATGCTGTATGAAAAAGAGAGAGTGTCTACGGG 3938
QY 7044 taacggccaacttaaaagaagacttgccttggcggaatgaagggcttcaactgtgtgctgaag 7103
|||||
Db 3939 TAAAGGCGCCACTTAAAAAGAGACTTGGCTTGGGGGATGAGGGGCTTCACTGTGGCTGAAG 3998
QY 7104 cagaagaagtagctcaaaaagctgcgggaatttaagccgggaagggcgagatttagttg 7163

```
Db 3999 CAGAGAAGTGTAGCTCAAAAAGCTGCGGAATTTAAACCGGAAGGGGCGAGATTTAGTTG 4058
Oy 7164 attcgaatacaacctgaagcctgaagaataaactccaattctctcccaagtca 7223
Db 4059 ATTGCGATACACACTAGCCTGGAAGGAATTAATCTCCAAATTTCTTCTCAAGGTCA 4118
Oy 7224 aggcgtgcggttccctcaagatgacggaaggaactgaagcgtgcggaat 7283
Db 4119 AGGGCTGCGCGTTTCCCTCAAGTATGCCGGTAGGGAATACTTGAAGCGGCTGCCGGAAT 4178
Oy 7284 ttacctctgcggcctctgcggaagcagcagcggcggtggttaagcgtatgctcgg 7343
Db 4179 TTACCTTGTGCGGGCCCTGACGGAACAGACGGGGAGCTGCTTACGGCTATGATCTGG 4238
Oy 7344 ccgggtggaataatgcggaagatgcccgttgaagaacatcagcgcggaagtcaag 7403
Db 4239 CCGGTTGGGAATAATCCCGCAAGTATGCCCGTTAAAGACATCAGCTGCGAGGTCAAG 4298
Oy 7404 aagactccgttttggcaagaatgcttaccgttgctcgtcgttaagggatgattaa 7463
Db 4299 AAGACTCCGTTTGGTCAAGACTGCTTACGTTGCTGCTGCTTAAAGGGTGATTTAA 4358
Oy 7464 ccgtgacctatgaagtgatgacgagcggaagatgctgctgaacagctgctccagcg 7523
Db 4359 CTGTGACCTATGAAGTGCATGAGACGGGGCAAGATTGCTGTACACGCTGACTCCAGGCG 4418
Oy 7524 cgggaagaagctggtctctctgcgaagccttgcttgaaccttgccctgcgaagaatga 7583
Db 4419 CGGAAGAAAGCGGCTCTCTGCGACGCTTGTGGCTTGAACCTGAGCCCTGCAAAAAGACTGA 4478
Oy 7584 ccgaataccgactactatgctctggaagcctaaatgaagatgaccagcgccttggaagta 7643
Db 4479 CCGATTACCGGTACTAGTGTCTGGGACTTAATGAGAGCTAACCAACCGCTTGGAAAGTA 4538
Oy 7644 attactggtgcatctacgaaggaagcgttaaaaaaagaacttaagcccaata---tcgtccg 7700
Db 4539 ATTACTGTGGGATCTAACGAGGAGCGGTAAAAAAGAACTTAAAGCCATACCTGCGCTCCGC 4598
Oy 7701 aggaagaagcgaagcggagcgaagctgctgctgacagctcttgatgaagaagcggct 7760
Db 4599 AGGAAGCGGCAACCGGAGCAAGGTTCCCTGTGACCACTCTTTTATGAAAAGGGCGGCT 4658
Oy 7761 tgaattacgccaatgaggggagcgaacttgactgctgcttgccatattctgcgcgc 7820
Db 4659 TGAATTATACGGCCATGCGGGAGACTTGAACCTGTGCTTGTCCATATCTGCGGCC 4718
Oy 7821 aaattgaagcagcggaacagcctttgaactgactaacaacttaacttgggttaagcct 7880
Db 4719 AAATTGAAGCAGCGGACACGCTTTTGAACGTGATTAACAAATTACACTTGGGTTAGAGCTT 4778
Oy 7881 taagcgcaccaatggggctcgccggggaatgactccttgagggagagaaggtccaccggaa 7940
Db 4779 TAAGCGCCCAATGGGGGTCCGGCGGGGATGACTCCTGGGGGAGAGAGTCCACCGGAA 4838
Oy 7941 tctgcctgagatgctcaaaaagccgcagccttgcctgctgattcaagcccttactaa 8000
Db 4839 TCTGCGCTGATGCTCAAAAAGCGCCGACGTCCGCTGCTGATTCAGCCCTTTTAACTAA 4898
Oy 8001 aataatgctcaatgacttaacaggaatgaatttaagtaaaagcgaagtgagga 8060
Db 4899 AATAAATCTCAATTTGACTTAACAGGATGAATTTTGTATTAACCAAAAGCAGTGAAGA 4958
Oy 8061 agatggcgaagatcagaagaatgccaagcagcggtgctcgtcagcgagcgt 8114
Db 4959 AGATGGCAGACATGAGAGAAGTGCACAGGACGCCGCTGTCTAGCGACGCT 5012

RESULT 3
LBALACZBUL 5015 bp DNA BCT 26-Apr-1993
LOCUS L.delbrueckii bulgaricus lactose permease and beta-galactosidase
DEFINITION (lacZ) genes, complete cds.
```

```
ACCESSION M55068 M80754
VERSION M55068.1 GI:149564
KEYWORDS beta-galactosidase; lactose permease.
SOURCE L.delbrueckii bulgaricus DNA.
ORGANISM Lactobacillus delbrueckii
Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillaceae;
Lactobacillus.
REFERENCE 1 (bases 1 to 5015)
AUTHORS Leong-Morgenstaler,P.M., Zwaalen,M.C. and Hottinger,H.
TITLE Lactose metabolism in Lactobacillus bulgaricus: Analysis of the
JOURNAL primary structure and expression of the genes involved
MEDLINE J. Bacteriol. 173, 1951-1957 (1991)
FEATURES
location/Qualifiers
1..5015
/organism="Lactobacillus delbrueckii"
/sub_species="bulgaricus"
/db_xref="taxon:1584"
/feature_1db="ATCC 11842"
102..1985
/gene="lactose permease"
102..1985
/gene="lactose permease"
/note="putative"
/codon_start=1
/transl_table=1
/product="lactose permease"
/db_xref="GI:149565"
/transl_table=1
/product="lactose permease"
/db_xref="GI:149565"
/translation="MKKKVLSRLSYAAGAGNDVYATLSTFYFVETTHLPFNADHK
MFIITNLITAIRIGEVLDLPJLGNADIDRESRWKRPVWVGGLISLALFLPD
FGFINGOSKPVVLYLFGIVLYLMDIFDSFKUTGPMWAI PALSDSREKRTSPARG
STIGANLVGVITPILIEFSASKAMPNDGKGWFFALIVAVIGILSITVGLTHER
KSALRESNEKTTLOVRFVGNOLMLAAYWYEGINTALDLYRSYTLGA
RGVSLITYTFVGLISASPPRSIAKPNRRLYACIAVWLIGIVSAVSGSLAS
LVGAEFFPTPLFLVYVLMITSAYEFGKLTGRBREALTSTRPYDKGLSWM
FVSLIALTAGMTTATATSTTAHOMYFKLAMEALPAVMLITAVSTPAKVFLEENH
AEIYDOLETOFGOASHAQ PAQAESEFTLASPVSGOLMDIWDVDDVFAKDKIGDFALV
PADGKVVYAPFAGTIVROLAKTRHSIVLENEHGLVLIHLGLTAKLNGTGFVSVEEGS
OVEAGQOILFEWMDPAIKQAKDDIVIVYIVINSETFANSOMLPIGHSVQALDVFKLE
GKN"
1989..5015
/gene="lacZ"
1989..5015
/gene="lacZ"
/note="putative"
/codon_start=1
/transl_table=1
/product="beta-galactosidase"
/protein_id="AA25244.1"
/db_xref="GI:149566"
/translation="MSNKLVEKRVDOADLMLTPDEYEVNTIPSHDHSFQSOEE
LEEGSSIVOSIDGDMLIDVAENGOGVNFYAEDFDPSNFSVAPGMLTELOGGGOO
YVNVQYPRDGSSELEFPQIIPSKNPLASVYRFTDDEAFMDKEVSLKDGATATYVM
NGHVFVGTGDSFTFSEFNVTKFLKENRLVALIKTSSASMLEDDQFWRKATGLFRY
TLOAKPRHLEEDLKTASLTDNYOKGLVEPANTAYRLPNASFLEVRDSEGLVAEK
LGPRISEOLEFTLADLPVAAMSAKPNLYQVRLTGLFEGLEPESGTFALMPBEVDEGQ
EVYVRVTAHLKEDLPMADEGTVVAEEVAGOKLPEFKEGEPDLVDSDVNLGLKNNF
OIEFSKYGMPVSLKYVAGRETKRLPEFTFRALTDMDRAGVGYDLARWENACKVYR
LKDI SCEKEDSVLYKTAFTIPLAKLEDIVTYTVEYDGRKIAVYADPDGAEAEALLPA
FGMLVALPEKLETFDRIYVGLPNESTYPRLEBNYLGITVQAGVKKNFSPYLRQEGNS
KVRYQLDEKRGLETFANGADLNLISALPYSAQIEADAHAFELTNNYTVWRALSAQM
GVGDDSDMGCVHPEPCLDOKARQLRVLOPLLK"
```

BASE COUNT
ORIGIN

1231 a 1322 c 1292 g 1170 t

Query Match		41.7%;	Score 3386.6;	DB 3;	Length 5015;
Best Local Similarity		99.5%;	Pred. No. 0;		
Matches 3408;		Conservative 0;	Mismatches 14;	Indels 3;	Gaps 1;
OY	4584	ccgataaaagttaagcgacgagcttgcgcttgcgagcagacggtlaagcttaacgagc	4643		
DB	1591	CCGACAAAAAGTTAGCGAGAGGCTTTGCCGTGGCCAGCAGACGGTAAGCTACGCCG	1650		
OY	4644	catttgcggtactgtccgcagctgccaagaccgagcactcgatcgttccctggaatg	4703		
DB	1651	CATTTCGCGGTACTGTCCGCGACCTGGCCAAAGCCGGCACTCGATCTCTGGAAATG	1710		
OY	4704	aaatagggtctgtgcttgcattacaccttgcgctggaagcagctcaaatlaaagagatg	4763		
DB	1711	AACATGGGCTCTTGCTTGTATTCCTTGCTGGCGACGGCAATTAAGGGGACTG	1770		
OY	4764	gcttttgcagctactgttgaagagggcagccaagtagaagccgagccaagatcctggaat	4823		
DB	1771	GCTTTGTCAGCTATGTTGAAGAGGGCAGCCAGGTAGAGCCGGCCAGCATCTGTGAAT	1830		
OY	4824	tctgggaaccgagcatcaagcaagccaagctggaagcagcagctgaatcgtagccgtaaca	4883		
DB	1831	TCGTGGAGCCCGGGATCAACGACGAGCTGGACGACGAGGTAAATCGTACCGTCAATCA	1890		
OY	4884	acagcgaaacttccacaataatagccagatgctcttgcgcatcgccaagcgttcaagccc	4943		
DB	1891	ACAGCCAAACTTTTCGCAATAGCCAGATGCTCTTGCCGATCGGCCACAGGCTCCAAAGCCC	1950		
OY	4944	tggatgatgtatcaagcttgaagaggaagatagaanaatgaagcaatlaagctlaagtaaaag	5003		
DB	1951	TGGATGATGTATTCAGATTGAGAGGGAGAAATTAGAAAAAGACCAATTAATTAGTAAAG	2010		
OY	5004	aaaaaagattgacagcagcagcagcctgctgctgctgactgaagccgggaagtttacgaagta	5063		
DB	2011	AAAAAAGATTGACACGAGGAGACCTGGCCCTGGCTGACTGACCCGGAGATTTCAGAGTCA	2070		
OY	5064	atacaattcccccgcaactccgacatgaagtccttccaaaacaggaagaaacttgaagagag	5123		
DB	2071	ATACAAATTCCTCCCGCACTCCGACCATGATGCTTCCAAAGCCAGAGAACTGGAGAGG	2130		
OY	5124	gcaagtcagctttagtgcagctcccttgcagcggagagctgctgactgaactacgctgaagag	5183		
DB	2131	GCAAGTCAGATTAGTGAAGTCTCTTGAGCGGGAGCTGGCTGATTAAGTACGCGAAAGCG	2190		
OY	5184	gccagggagcagctcaactctcatgcsagaagactttgaagatagcaattlaagtaagta	5243		
DB	2191	GCCAGGAGACAGTCAACTTCTATGACAGAAAGACTTGTGACGATACCAATTAACTTCAGTCA	2250		
OY	5244	aagtacccggcaacctggaactgcaaggtttggccaagcccgactatgtaaaagttcaaat	5303		
DB	2251	AAGTACCCGGCAACTGGAACCTCAAGGCTTTGGCCAGCCCAAGTATGTAACGTCCAAT	2310		
OY	5304	atccatggagacggcagtgaggaagatttcccgccccaattccaagcaaaaatccgctcg	5363		
DB	2311	ATCCATGGGAGCGGCACTGAGAGATTTTCCCGGCCCAAAATTCACAGAAAAATCCCGTCG	2370		
OY	5364	ctctctaatgtcagaatacttgcagctggaagagcttctgaggacaaggaatgaagcttga	5423		
DB	2371	CTTCTTATGTCAATACTTGTGACCTGGATGAACTTCTTGCGAAGGAAGTAGTACGCTTGA	2430		
OY	5424	agtttgaagggggcggaacagcaataatgtctgctgctgagcagccactgtgtgctgag	5483		
DB	2431	AGTTTACGGGGGCGCAACGCCATCTATGTCGTGGCTGAAGCGCCACTTGTGTGGCTACG	2490		
OY	5484	gggaagaactccttaccaccaagcgaatttlaagttaccaagcttccccaagaagaaataata	5543		
DB	2491	GGGAAGACTCTTTACCCCAAGGAGTTTATGTTACCAAGTTCTCTCAAGAAAAAACA	2550		
OY	5544	accgcttgagcagtgctgtctctacaagatattcttcgacctctctgctgctggaagcagact	5603		
DB	2551	ACCGCTTGAGAGTGCTCTCTACAAATATTCTTCCGCTTCCTGGCTGGAGACACGACT	2610		
OY	5604	tctgggcacatgctgtgttgcagatcaagtgactcttcaagcgccaagccgctgctgact	5663		

DB	2611	TCGTGGCCATGTCTGTGTTTTCAGATCACTGACTCTTTCAGGCCCAAGCCGCTGTGACT	2670		
OY	5664	tgaagaccttgaagcttacgcgagcagcttgaacgataactacccaagaaagaaagctggaag	5723		
DB	2671	TGAGAGCACTTAAGCTTACGCGCAGCTTGACCCGTAATCTCAAAAAAGAAACCTGGAG	2730		
OY	5724	tgaagccaatattgtcctaccgcttgcgcaaatgccaagcttlaagcttgaagctgcggata	5783		
DB	2731	TCGAAGCAATATTGCTCTTACCGCTTGCCAAATGCCACCTTTAAGCTGGAAGTGGGATA	2790		
OY	5784	gtgaaggtgacttggcttgcgaagaagcttggcccaatcagaagcagcagctgaattca	5843		
DB	2791	GTAAGGTGACTTGTTGCTTGCTGAAAAGCTGGGCCCAATCAAGCAGCACTGGAATTCA	2850		
OY	5844	ctctgctgacttgcgcaagtagctgctctggaagcgcggaagaaacctaacttaccagttcc	5903		
DB	2851	CTCTGGCTGATTTGCCAGTAGCTGCTCGAGCGGAAAGCCTAACCTTTACAGGTTCC	2910		
OY	5904	gctctgtattataccagggcagcagcctcttgaaggttgaagctgcggcaggaagtgttcc	5963		
DB	2911	GCTGTATTTATACACGACGACGACCTCTTAAAGTTACCCGGCAGGAAGTGGGTTCC	2970		
OY	5964	gcaacttgaactaaagaagcggatlaatgaacttgaacgccaagcagctgcttcaag	6023		
DB	2971	GCAACTTTGAACATAAAGAGCGGATTTATGACTTAAAGCCGACGAGATGCTTCAAG	3030		
OY	6024	gggccaaccgacgaaatttgaacagtaagtgtggtcgggtatcaacggaagagata	6083		
DB	3031	GGGCAACCGGACGAAATTTGACAGTAAAGTTGGGCCGGCTATCACGAAGAGATATGA	3090		
OY	6084	tctgggaacatcaagaacatgaagcgaagcaaatcaatgctgtcgtcttccactacc	6143		
DB	3091	TCGTGGTATCAAGACCATCAAGCAGAACTCAATGCTGCTCCGCTCACTCAAC	3150		
OY	6144	cgagcagctccctcttcttaccgctctgtgacaaagtaacgaccttgaatgaatgaag	6203		
DB	3151	GGAACCACTCCCTCTTTTACCGGCTCTGTGACAAAGTACGGCCTTTAGCTATGATGAAG	3210		
OY	6204	ctaaccttgaagaaacacagcagccttgggaagaaagtgtgggggcaagaaagctcactga	6263		
DB	3211	CTAACCTGGAAGACGACGCACTGGGAAAAAGTGGGGGGGCAAGAAATCTTAGCTTCA	3270		
OY	6264	atgttccagggcagatgacacagcatgtgctggaagccagcttataccgggtgaagaaatga	6323		
DB	3271	ATGTTCCAGCGCATGACACCACTTGGCTGGAGCCACACTTATCCCGGTAAACATGA	3330		
OY	6324	tggctcggaagcaagaacacatgcttcaatccttaacttgccttgaaggaatgaatgcttaag	6383		
DB	3331	TGGCTCGGACAAAGAACCATGCTTCAATCTTGATCTGGCTTTTAAAGCAATGAGCTTACG	3390		
OY	6384	ccggcaactgtcttcttgcgcaaatgctgctgattacgttccggaagcttgaaccgagttc	6443		
DB	3391	CCGGCACTGTCTTGGCCAAATGGCTGATTAAGTCCGGAAGGCTGATCCGACCCGGGTTG	3450		
OY	6444	agcaactatgaaggggtgaaccacaacccggaagtttgaagcgcgaaccagatgtgaagcc	6503		
DB	3451	AGCACTATGAAGGGGTGACCCACACACGGAAGTTTGAAGCAGCGCACCAATGGAAGGCC	3510		
OY	6504	ggaatgatagtccggccaaggttaattgaagaatacttgaaccaataaacaggccaagcat	6563		
DB	3511	GGAATGATGCTCCGGCCAAAGGTAAATTGAACAAATCTTGACCAATTAACGCCAAGCAT	3570		
OY	6564	ttaactgaattgaatacgtctcacgcagctatgggaactccgctgcttgaacttgcgctcaca	6623		
DB	3571	TTATATCAGTTGAATAGCTACAGCCATGGGCAATCCGCGGAGACCTGGCCGCTACA	3630		
OY	6624	cggccttggaagaaataaccccactacacagggcgctctcaatctggaatgtgaattgaacaa	6683		
DB	3631	CGGCGCTGGAAGAAATACCCCACTACAGGGCGGCTTCACTGGAGATGGAATTGACCAAG	3690		
OY	6684	gaactggaagaaagcggcagctgtcttgaaggcgactcgaatgacggccaagcagact	6743		

Db 3691 GACGTGAAAAAGACGGGACCTGCTTTATGSGGGGAGCTTGCGGAGACCGCCCAACCGACT 3750
Qy 6744 atgaattcgcgggaacgcgcgtgtcttctgtaccgcgcgaatcgcgcgaactcgtgta 6803
Db 3751 ATGAATTTGGGGGAGACGGCTGTGCTTGTGACCGGACGTAATCGCCGAACTGGGCTA 3810
Qy 6804 atgtcaaggcccttaacgcgaacttaagttagaagtaaaagatgaggcagctcttccta 6863
Db 3811 ATGTCAAGGCGCTTACCCCAACCTTAAGTTAGAAGTAAAGATGGGACAGCTCTCTCA 3870
Qy 6864 aaacgcgaattttttaccacagctcatcttactacttcttgactgtcttctgtgcg 6923
Db 3871 AAAACGAATTTATTTATCAACAGCTCATCTTACTCTTGTGACTGTGCTTTTGGTGG 3930
Qy 6924 atggcaattacacacagagccgcgcctcgcaccttggccttgagccttgccgaatcg 6983
Db 3931 ATGGCAATTACCTACACAGAGCCGGCTGTGACCTTTGGCTTGGAGCCTGGCGAATCG 3990
Qy 6984 ggaaccttgccctgccttgccgcgaagtcgctgatatgaaaaaagggaggctgctacccgg 7043
Db 3991 GGACCTTTGGCCCTGTGCGCGGAATCGCTGATGAATAAAGAGAGGTCGTCTACCGGG 4050
Qy 7044 taagcgcccaattaaaaaagaacttgcttgaggagtgagggcttcactgtgctgaag 7103
Db 4051 TTACGGCCCACTTAAAAAAGAACTTGCCTTGGCGGATGAGGGCTTCACTGTGGCTGAAG 4110
Qy 7104 cagaagaagtgcacaaagctgcgcgaatttaagccgaagggcgccgacagattagtg 7163
Db 4111 CAGAAAGAAAGTACCTCAAAAGCTGCCGGAATTTAACCCGAAAGGGCGCCAGATTAGTTG 4170
Qy 7164 atccgcatacaactagccttgaaagaaataacttccaaattctcttccaaagtca 7223
Db 4171 ATTCGCACTACAACTAGGCTGGAAGGAATACTTCAAAATCTCTTCTCCAGAGTCA 4230
Qy 7224 agggctgcgcgttcttccctaagatagtcgcgtaaggaaatacttgaaagcgctgcggat 7283
Db 4231 AGGGCTGGCCGGTTTCCCTCAAGATAGCCGATAGGGAATCTTGAAGCGCGCTGCCGGAAT 4290
Qy 7284 ttactcttcgcgcgcgccttcgcgaacacgcgcggagcggttaagggcatgctcgg 7343
Db 4291 TTACTCTTGGCGGGCCTTGACGAGCAACGACCGGGAGCTGTGACGGCTATGATCTGG 4350
Qy 7344 ccgcgtggaagaaatgcgcgcgaagatagccgccttgaaagacatcagctgcgagtgcaag 7403
Db 4351 CCCGGTGGGAAATAGCCGCGCAAGTATGCCGCTTGAAGACATCGTGCAGGTCAGAG 4410
Qy 7404 aagctccgcttggtaagagctgccttaagttgctgcctgccttaaggtgattaa 7463
Db 4411 AAGACTCCGTTTGGTCAAGACTGCTTACGTTGCTGCTTAAAGGCTGATTTAA 4470
Qy 7464 ccgcgtacatgaagtcagtgagcgaggaagatgctgtaacagctacttccagcg 7523
Db 4471 CCGGACTGATGAAGTGCATGAGCGGGCAAGATTGCTGTAAACAGCTACTTCCAGGCG 4530
Qy 7524 cggagaagatcgtctcttgcacgccttggcttgaaacctgcgcgcgcgcgcgaagactga 7583
Db 4531 CGGAAGAAAGCGGCTCTTGGCACGCTTGGCTTGAACCTAGCCCTGCCAAGAACTGA 4590
Qy 7584 ccgattccgctactatgcttgtagactaagtagagcttacccagaccgccttggaagta 7643
Db 4591 CCGATTACCGCTACTATGCTGTGGACCTAATGAGCTTACCCCAACCGCTTGGAGAGTA 4650
Qy 7644 ataccctggagcatctacagaggaagcggttaaaaaaagaactttagccacata---tcgtccg 7700
Db 4651 ATTACTGTGGCATCTACAGGAGCGGTAATAAAGAACTTTAGCCCATTAAGCTGCGTCCGC 4710
Qy 7701 aggaacgaggaacccggaagatctgcgtgtaccagctcttgaatgaagggcggt 7760
Db 4711 AGGAAGCGGCAACCGGAGCAAGGTTGCTGTACACAGCTCTTGTGATGAATAAAGGCGCT 4770
Qy 7761 tggaaattacggcaatgagggaagactgaaactgtctgtcttgccatatctcgcgcc 7820
Db 4771 TGGAAATTACGGCAATGGGGGAGACTTGAAGCTGTGCTTGGCATATTCTGGCGGCC 4830

Qy 7821 aaattgaagcagcggaacacgcctttagaactgactgaactaacaattacacttgggtttagacct 7880
Db 4831 AAATTGAAGCAGCGGACACCGCTTTTGAACGTAGCTAACAATTAACCTTGGGTTAGAGCT 4890
Qy 7881 taagcgcgaatgagggttcggcgaggatgactcctgggggcagaagttccaccggat 7940
Db 4891 TAAGCGCCAGATGGGGGTGCGCGGGAGTGAATCTCTGGGGGCGAGAAGGTCCACCCGGAAT 4950
Qy 7941 tctgccttgatgctcaaaaagccgcagcttgccttgatgactgaagcccttactaa 8000
Db 4951 TCTGCTGATGCTCAAAAAGCCGCGACGCTCGCTGTGATTAAGCCCTTTACTAA 5010
Qy 8001 aataa 8005
Db 5011 AATAA 5015

RESULT 4
EVEI32038
LOCUS EVEI32038 3840 bp mRNA circular SYN 28-JUL-1999
DEFINITION Expression vector pCDPr.
ACCESSION AJ132038
VERSION AJ132038.1 GI:5640088
KEYWORDS AMP gene; beta lactamase; ColE1 origin of replication; multiple cloning site; Sp6 promoter; SV40 origin of replication; T7 promoter; xanthine-guanine phosphoribosyl transferase; Xanthine-guanine phosphoribosyl transferase gene.

SOURCE Expression vector pCDPr.
ORGANISM Expression vector pCDPr.
REFERENCE 1 (bases 1 to 3840)
AUTHORS Zeng,B.J.
TITLE Mammalian Expression Vector for with fuse Xanthine-guanine phosphoribosyl transferase Tag
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 3840)
AUTHORS Zeng,B.J.
TITLE Direct Submmission
JOURNAL Submitted (27-FEB-1999) Zeng B.J., Gene Engineering Center, Institute of Microbiology, Zhongguancun, Beijing, Beijing 100080, CHINA

FEATURES
source Location/Qualifiers
1..3840
/organism="Expression vector pCDPr"
/db_xref="taxon:90749"
209..863
/note="CMV"
864..882
/note="T7"
882..984
/note="Multiple cloning site; HindIII, BamHI, BstXI, EcoRI, NotI, XhoI"
929..1387
/codon_start=1
/product="Xanthine-guanine phosphoribosyl transferase"
/protein_id="CAB51567.1"
/db_xref="GI:5640089"
/translation="MSKKYIVTMDMLQIHARKLASRMPSEOWKGIIVSRGIVPGA LPRAEIVRTIFAKRGRPLVDVYDIPDTLWLEQPMWDMGVFVPPISGR"
1649..1863
/note="BGH"
/note="Sps"
2450..2775
/note="Sps6"
2644..2729
/note="SV40"
/complement(2844..3704)
/gene="amp"
complement(2844..3704)
/gene="amp"
/codon_start=1
/product="beta-lactamase"

[illegible][illegible]


```

repeat_L_region
1335..1351
/Note="direct repeat A"
repeat_L_region
1352..1368
/Note="direct repeat B"
repeat_L_region
1369..1385
/Note="direct repeat C"
repeat_L_region
1386..1402
/Note="direct repeat D"
repeat_L_region
1403..1419
/Note="direct repeat E"
repeat_L_region
1420..1436
/Note="direct repeat F"
repeat_L_region
1437..1453
/Note="direct repeat G"
repeat_L_region
1456..1467
/Note="direct repeat H"
repeat_L_region
1468..1484
/Note="direct repeat I"
repeat_L_region
1485..1501
/Note="direct repeat J"
repeat_L_region
1504..1520
/Note="direct repeat K"
repeat_L_region
1521..1537
/Note="direct repeat L"
repeat_L_region
1538..1554
/Note="direct repeat M"
repeat_L_region
2051..2064
/Note="origin of replication (put.): putative"
BASE COUNT 639 a 335 c 466 g 653 t
ORIGIN 1 bp upstream of HindIII site.

```

```

Query Match 15.3%; Score 1238.4; DB 3; Length 2093;
Best Local Similarity 98.0%; Pred. No. 2.6e-285;
Matches 1254; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

```

```

QY 2032 gacacattcccccgaagagcgacacgagcgagcgagagatcatcctgacat 2091
DB 1280 GCACGCTGCACAAACAAAGGATATGCCACGAGATTTACCGTGTATTCCTACAT 1221
QY 2092 tctcttaccacaataaatttggtttatataaataccatttggcgaaactctcc 2151
DB 1220 TCTCTTACCAATATAAATTTTGTATTAATCCCATTTTGCACACTTCTTCC 1161
QY 2152 gaagcttcatttgccttcttggtaattcattcgcaacatgaaataacacatc 2211
DB 1160 GCAGCTTCATTGCTTTGGTATATCTTCATCGCAACATGAATTAATCACCATT 1101
QY 2212 tcaaatcttcaagttcaaatcttggcttaatttggcttaataacacatagctgat 2271
DB 1100 TCAACATCTTCAAGTTTCAATCTTCTTAATTTGCTTAATATACACATAGCTGATT 1041
QY 2272 tgcgtgtctcagcgaagacatactccaaatttlaatacaccacaatacagctcatca 2331
DB 1040 TGCAGTTCAGGCTTAAGCATATCTCAATTTTATATACGCCAAATTAAGCCTCATCA 981
QY 2332 tcaagcgtcatataatcagctgatttcttactcgttattgcgcttcttcggaactagct 2391
DB 980 TCAGCGCTCATATTAACACTGATTTTACCTGATTTTGCCTTTCTTCGGACATAGCT 921
QY 2392 tgcgaagagtcagcttcttacttgcgttgaagcttgaagcagcgcagcagcagcagc 2451
DB 920 TGCAGAAAGTCAAGTCTTACGTTTGTATGCTTTAACAACCTGCAATGACACCAAGGC 861
QY 2452 tcaataacacttcaagcgttcttgcataaatttgcacatctgcgttgcgttaataa 2511
DB 860 TCATATATCAACTTCAAGGCTTTTGCATATTTTGCACATCTCTGCTGTAAATTA 801
QY 2512 ttatttgaattctttaaataactgatttcaacaagcaacatgacagtgattat 2571

```

```

DB 800 TTATTGAATTTCTTAAATTAACCTGATTTTACAAACACACATGCAAGTGTGATTA 741
QY 2572 tatgacccgtctgtctcatlaacgtaattcccgcttgaacgtaataaocccataaatt 2631
DB 740 TATGACCCGCTGTGTTTCAATTAAGGTAATTTCCGTTGAACGTAATTAACCCATTAATT 681
QY 2632 ttaagtaactttttatagagatgacttataaaggctttagtaagcttcttaagac 2691
DB 680 TTAGTACACTTTTATATAGGAGTATGCTTAATTAAGGCTTTAAGTCAAGCTTTAAAGAC 621
QY 2692 actttaactctctgcgtgaatgagcgtttttaagcgtttaaagcttaaaacaacacgt 2751
DB 620 ACTTTTAACCTCTCTGCTGTAATGACGTTTAAAGGTTAAAGTTAAACCAAAACCGT 561
QY 2752 cctttagcctctcttgcacactgtctcgcgaataatttcttaactgctcagatttctc 2811
DB 560 CCTTTAGGCTCTCTTGAACACTGCTCCGCAATATTTGTTTAAGTGGCTGCAAGTTTTC 501
QY 2812 atgtcctcttccaaattacacaaatgagacaaatcglttatgacaaacacgcttgatata 2871
DB 500 ATGCTCTCTTCCCAATTAACAAATGACACAAATCGTTTATGACAAACACGCTTTGATTA 441
QY 2872 agttttaagtgctgcgaactcttaacgaacgaacgaacacttcaacccgatacatca 2931
DB 440 AGTTTAAAGTGTCTCGCCCAATCTTACGAAACCCCAAACTTCAACCAACCCGTAATCA 381
QY 2932 tgtcccgcttlaaattcttaagattgccaataatctgcgaatagcgaacatttcaactc 2991
DB 380 TGTCCCGCTTAAATTTCTAAGATTTGCCAATATTTGCGCATACGCAATTTCAATCTTTC 321
QY 2992 cgtctgcgaaggtcctaacttgcatttcaatttcaatttcaaaaaattcttgcata 3051
DB 320 CGTTCGCGCAAGGCTTAACCTTTCGCAATTTTCAAGTTTATTCGCAAAATTTCTGACATA 261
QY 3052 aaaaagctctcgaatttcaacacgtgaagagcgtactcttcttcaataagcttataa 3111
DB 260 AAAAGCTCTCCAGTTTATTCACGTAAGAGACTGACTATCTTTTCATTAAGCTTATTA 201
QY 3112 ccttgcacatcagagcttccctcctagataagctataaactgcgaactgataaact 3171
DB 200 CCTTGACATCATATAGGCTTTTCCCTAGATTAAGGCTTAATTAATGCAAAATTAATCACT 141
QY 3172 caagtgcttcgagcgcgaactaggaatttgcagtggtgtttatttgccttctctt 3231
DB 140 CACGCTTTCGAGCGGCCCAACTAGGAATTTGCACGTGGCTTTTATTTTGTCTTCTT 81
QY 3232 caaccaatttaaaccttaataacacaaagcctataaaatcaatgatacaagccc 3291
DB 80 CACGCAATTTATTAACCTTAATATACACCAAAAGCTTATTAATCAATGATTAACAAGCC 21
QY 3292 aattaagcctaatacaagctt 3311
DB 20 AATTAGCCTAATCAAGCTT 1

```

```

RESULT 7
C300RFR 2140 bp DNA BCT 07-MAR-1996
LOCUS Plasmid pc3011 (from L. plantarum) ORF R, complete genome.
DEFINITION J03319
ACCESSION J03319.1 GI:144180
VERSION complete genome.
KEYWORDS plasmid pc3011 DNA.
SOURCE plasmid Plasmid pc3011
ORGANISM plasmids.
REFERENCE 1 (bases 1 to 2140)
AUTHORS Skaugen, M.
TITLE The complete nucleotide sequence of a small cryptic plasmid from Lactobacillus plantarum
JOURNAL plasmid 22 (2), 175-179 (1989)
MEDLINE 90160873
COMMENT Draft entry and computer-readable sequence for [1] kindly submitted by M.Skaugen, 19-NOV-1989.

```


JOURNAL Unpublished
REFERENCE 2 (bases 1 to 3986)
AUTHORS Peters, H.
TITLE Direct Submission
JOURNAL Submitted (07-AUG-1995) H. Peters, Inst. f. Immunologie,
Michaelisstr. 5, D-24105 Kiel, FRG
COMMENT Related sequences: M21295 and K03104.
FEATURES
Location/Qualifiers
source 1..3986
/organism="synthetic construct"
/plasmid="pcDNA3.2EO"
/db_xref="taxon:32630"
misc_feature 1..2125
/note="cloning vector (pcDNA3) (Invitrogen)"
misc_feature 889..394
/note="multiple cloning site (MCS)"
misc_feature 2126..2796
/note="cloning vector (pZeoSV) (Invitrogen)"
misc_feature 2797..3986
/note="cloning vector (pcDNA3)"
BASE COUNT 900 a 1057 c 1054 g 975 t
ORIGIN
Query Match 13.3%; Score 1082; DB 56; Length 3986;
Best Local Similarity 100.0%; Pred. No. 7.3e-248;
Matches 1082; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 gatgaggggccaagatagcgggtgacattgattatgacagttatataagaaatca 60
Db 210 gatgtagcggggccagatgatacggcttgacattgattgacagttatataagaaatca 269
QY 61 attacggggatcatagttacatagccatataatgagatccgcgtacataactacgta 120
Db 270 attacggggatcatagttacatagccatataatgagatccgcgtacataactacgta 329
QY 121 aatggcccgctgctgacacggccaagaccccgccatgagtcgaataagcgtat 180
Db 330 aatggcccgctgctgacacggccaagaccccgccatgagtcgaataatgacgtat 389
QY 181 gtcccatagtaagcccaatagggacttcattagtcgaataagtgagacttaagc 240
Db 390 gtcccatagtaagcccaatagggacttcattagtcgaataagtgagacttaagc 449
QY 241 taactgcccacttgacagatcaatgtaataatgacaaagacacccctatgac 300
Db 450 taaactgcccacttgacagatcaatgtaataatgacaaagacacccctatgac 509
QY 301 gtcaatgacgttaaatgacccgctgacattagccagataagacattatggactt 360
Db 510 gtcaatgacgttaaatgacccgctgacattagccagataagacattatggactt 569
QY 361 cctactgagcagatatactagctatgatacgtatcattacatgtagtggtttgg 420
Db 570 cctactgagcagatatactagctatgatacgtatcattacatgtagtggtttgg 629
QY 421 cagtaacataatgagcgtgagatagcgtttgactcaacgaggttcccaagttccacccc 480
Db 630 cagtaacataatgagcgtgagatagcgtttgactcaacgaggttcccaagttccacccc 689
QY 481 attgacgtcaatbgaggtttgtttgacacaaatacaacggagacttcccaaatgctgt 540
Db 690 attgacgtcaatbgaggtttgtttgacacaaatacaacggagacttcccaaatgctgt 749
QY 541 aacaactcggcccatgacgaaatggcggtgagcgtgtacggtggaggttatata 600
Db 750 aacaactcggcccatgacgaaatggcggtgagcgtgtacggtggaggttatata 809
QY 601 agcagagctcctgtaactagagaacccactgttactgagcttaagaaatlaacg 660
Db 810 agcagagctcctgtaactagagaacccactgttactgagcttaagaaatlaacg 869
QY 661 actcaataagagaacccaagctgtgacgagctcgatccactagtaacgacgcca 720

|||||
Db 870 ACTCACTATAGGAGAACCAACCTGTGATACGAGCTCGGATCCACTAGTAACGCCGCCA 929
QY 721 gtgtgctggaattctgcagatatccatacaactgacgcccgttgagatgattcaag 780
Db 930 gtgtgctggaattctgcagatatccatacaactgacgcccgttgagatgattcaag 989
QY 781 ggcctatctatagtgtcaactaatgctagagctcgtatcagcctgactgct 840
Db 990 ggcctatctatagtgtcaactaatgctagagctcgtatcagcctgactgct 1049
QY 841 tctagttgcagacatcgtgtgtttgcccctcccccgtgcttctgacccctgaaagt 900
Db 1050 tctagttgcagacatcgtgtgtttgcccctcccccgtgcttctgacccctgaaagt 1109
QY 901 gccactccacgtctcttcccaataaataagagaattgtagtcattgtctgtagg 960
Db 1110 gccactccacgtctcttcccaataaataagagaattgtagtcattgtctgtagg 1169
QY 961 tctatctatctgtgaggtgtgaggtgagcagacgaagggagagattggaagac 1020
Db 1170 tctatctatctgtgaggtgtgaggtgagcagacgaagggagagattggaagac 1229
QY 1021 aatagcagcattgctgggagatgcggttggtctctatgcttctgagcggaagacagc 1080
Db 1230 aatagcagcattgctgggagatgcggttggtctctatgcttctgagcggaagacagc 1289
QY 1081 tg 1082
Db 1290 tg 1291
RESULT 9
AR098191
LOCUS AR098191 4026 bp DNA PAT 14-FEB-2001
DEFINITION Sequence 19 from patent US 6074850.
ACCESSION AR098191
VERSION AR098191.1 GI:12807448
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 4026)
AUTHORS Antelman, D., Gregory, R. J. and Willis, K. N.
TITLE Retinoblastoma fusion polypeptides
JOURNAL Patent: US 6074850-A, 19 13-JUN-2000;
FEATURES
Location/Qualifiers
source 1..4026
/organism="unknown"
BASE COUNT 978 a 1020 c 983 g 1045 t
ORIGIN
Query Match 13.2%; Score 1068.6; DB 9; Length 4026;
Best Local Similarity 69.7%; Pred. No. 1.2e-244;
Matches 1918; Conservative 0; Mismatches 89; Indels 746; Gaps 5;
QY 1 gatgaggggccaagatagcgggtgacattgattatgactagttatataagaaatca 60
Db 210 gatgtagcggggccagatgatacggcttgacattgattgacagttatataagaaatca 269
QY 61 attacggggatcatagttacatagccatataatgagatccgcgtacataactacgta 120
Db 270 attacggggatcatagttacatagccatataatgagatccgcgtacataactacgta 329
QY 121 aatggcccgctgctgacacggccaagaccccgccatgagtcgaataatgacgtat 180
Db 330 aatggcccgctgctgacacggccaagaccccgccatgagtcgaataatgacgtat 389
QY 181 gtcccatagtaagcccaatagggacttccattgagcgtcaatgggtgagactatgaag 240
Db 390 gtcccatagtaagcccaatagggacttccattgagcgtcaatgggtgagactatgaag 449

```
QY 241 taaatgccacttggcagtaacatcaagtgtatcatatgtccaagtacccccatttgac 300
|||||
Db 450 TAAACTGCCACTTGGCAGTATCATCAAGTGTATATATATGCAAGTACGCCCTATTGAC 509
QY 301 gtcaatgaagtaaatgccccgcctggcattatgcccagtaacatgaccttattggacctt 360
|||||
Db 510 GTCAATGACGGTAAATGAGCCCGCTGGCATATATGCCAGTACATGACCTTATGGAGCTTT 569
QY 361 cctacttggcagtaacatcaagtattagtaacgtatataacatgtgtatgagtggtttgg 420
|||||
Db 570 CTTCTTGGCAGTACATCTAGTATGTATCTATTCCTATTTACCATGTGTATCCGGTTTTGG 629
QY 421 cagtaacatgaatggcgttggatagcgtttgactacagggagattccaagtccacccc 480
|||||
Db 630 CAGTACATCAATGGGCGGTGTAGCGGTTGTACTACAGGGATTTCCAGTCTCCACCCC 689
QY 481 attgacgtcaatgggagttgttttttggcaccacaatcaacgggagcttccaaatgtcgt 540
|||||
Db 690 ATTGACGTCATGGGAGTGTGTTTGGACCAAAATCAAGGGAGCTTCCAAATGTGCT 749
QY 541 aacaactccgccccatgtagcgaatggcggttagcggttgaagtggtgagtgctata 600
|||||
Db 750 AACAACTCCGCCCATGTAGCCAAATGGCGGTAGCGGTGACGTGAGCTATATA 809
QY 601 agcagagctctgtgtaactagagaaccactgacttactgtgctatcgaatataacg 660
|||||
Db 810 AGCAGAGCTCTGTGGCTTACTAGAAACCCACTCTTACTGTGCTTATCAGAAATTAATACG 869
QY 661 actactataggagaaacccaagtctgg----- 687
|||||
Db 870 ACTGACTATAGGAGACCCAAAGCTTCGGGGGGGTACCACTCTCTCCGATCGGTGTG 929
QY 688 ----- 687
Db 930 CGAGGCCAGCTGTTGGGCTCGCGTTGAGAGCAAACTCTTCGGGTTCTTCAGTACTC 989
QY 688 ----- 687
Db 990 TTGATGGGAAACCCGTCGGCTCCGAACGCTACTCCGCCACCGAGGAGCCTGAGCAGT 1049
QY 688 ----- 687
Db 1050 CCGCATGACCGGATCGGAAAACTCTCGAGAACTGAAAAACGAAAGTTAACTGTGTA 1109
QY 688 ----- 687
QY 688 ----- 687
Db 1110 AGTTAGTCTTTTGTCTTTTATTTATTTAGGTCCCGATCCGGTGTGTGCAAAATCAAG 1169
QY 724 tgctggaattctgagatat----- 743
|||||
Db 1170 AACTGCTCTCAGTGTGATGTTGCCCTTACTCTTAGGCCCTGACGAACTGTTACTTCTGC 1229
QY 744 ----- 744
Db 1230 TCTAAAAGCTGCGGAATTTGTAACCGCGGCCCTGCAGTCTAGAGAAATTCGCTGATGATA 1289
QY 775 ctgaaaggccctattctatagtgtaacctaaatgctagagctgcgtatcagcctgcagt 834
|||||
Db 1290 TCGATGGGCCCTATTCTATAGTGTACCTAAATGCTAGAGCTCGCTGATCAGCCTCGACT 1349
QY 835 gtgctcttagttgcaagcaatcgtgtgttggccccctccccgtgcttcttcaacctg 894
|||||
Db 1350 GTGGCTTCTAGTTGCGCAGCCATCTGTTGTGCCCCCTCCCGGCTCTTCCCTTACCCCTG 1409
QY 895 gaaggtgcacctccactgtccttctcctaataaatgagaaatgtgcatgtcgtctg 954
|||||
Db 1410 GAAAGTGCCTCCACTGTCTTCTTAATAAATGAGAAATGTGCATTCGCTATGTCTG 1469
QY 955 agtagtgtaattctatctctgtaggggtgtaggttgggagcaagcaagggaggagattgg 1014
|||||
Db 1470 AGTAGGTGTCATCTATTTCTGAGGGGTGGGGTGGGCGAGACCAAGGGAGGAGATTGG 1529
QY 1015 gaagacaataagcaggtcgtcgggga----- 1040
|||||

Db 1530 GAAGCATATACCGAAATGACCGACCAAGGACGCCCAAGCTGCCATCAGAGATTTTGA 1589
QY 1041 ----- 1040
Db 1590 TTCCACCGCCGCTTCTATGAAGAGTTGGGCTTCGAAATCGTTTCCGGAGCGCGCTG 1649
QY 1041 ----- 1040
Db 1650 GATGATCTCCAGCGGGGGATCTCATGCTGAGTTCTTCCGCCACCCCAACTTGTATTAT 1709
QY 1041 ----- 1040
Db 1710 TGCAGCTTATATGTTTCAATAATAAGCAATAGCATCACAATTTCCAAATTAAGCAT 1769
QY 1041 ----- 1040
Db 1770 TTTTCACTGATCTTACTGTTGTGTGTTGCCAACTCATGATGATCTATCATGTCTG 1829
QY 1041 ----- 1040
Db 1830 TATACCGTCGACCTCTAGCTAGAGCTTGGCGTAATCATGTCATAGCTTTCTGTGTG 1889
QY 1041 ----- 1040
Db 1890 AAATTTTATCCGCTCACAAATTCACACACATACGAGCCGGAAGCATAAAGTGTAAGC 1949
QY 1041 ----- 1040
QY 1041 ----- 1040
Db 1950 CTGGGTCCTAATGAGTAGCTAATCTACATTAATTTCTGTTGGCTCAGTCCGCTTT 2009
QY 1063 ----- 1114
Db 2010 CCAATCGGGAACCTGTGCTCCAGCTCAGCTAATTAATGATCGCCAAACCGCGGGAGAGG 2069
QY 1115 cgtttgtgtaattggcgctcttcgcctctcgtactaactgactcgtcgtcgtcgtcgt 1174
Db 2070 CGGTTTGGCTATGGGCGCTTTCGGCTTCTCGCTCCTGCTACTGCTGCGCTCGCTGCT 2129
QY 1175 tcggtcgtcgagcgtgtaacgtactcaactcaagcggttaatacgttatacacaagaatc 1234
Db 2130 TCGGCTGGCGGAGCGGATACAGCTCACTCAAAAGCGGTAATAGCTTATCCACAGATTC 2189
QY 1235 aggggataacgcaagaaagaacatgttgaagcaaaagcgcaaaagcgcaagaaacgttaa 1294
Db 2190 AGGGATTAACCCAGAAAGAACATGTGAGCAAAAGCGCAGCAAAAGGACAGAACCCGTAA 2249
QY 1295 aaaggcgcgttctgtggttttccataggtctcgcgcgcctctgaagcaatcaaaaaa 1354
Db 2250 AAAGCGCGGTTGCTGGGCTTTTTCATAGGCTTCGCCCTCGACGATCACAAAAA 2309
QY 1355 tcgacgtcgaagtcagagtgtagcgaaaacccagcagaaataaagatacagagcgtttcc 1414
Db 2310 TCGAGCGTCAAGTACAGAGTGGCGAAACCGACAGAGATTAAGATATCAGGGGTTTCC 2369
QY 1415 cccctggaagctccctcgttgcgtctcgttgcgaacctgcgcgttaccggaatacgttc 1474
Db 2370 CCTGTGAAGCTCCCTCGGCTCTCTCTCCATCCACCTGCGGCTTACCGGATACGTGTC 2429
QY 1475 cgccttctcccttcgggaagcgttggcgcttccaaatgctcaagcgttgaagtatctag 1534
Db 2430 CGCCTTCTCCCTTCCGGAAACGTGGCGCTTCTCAATAGCTCACGCTGTAGGTATCTGAG 2489
QY 1535 ttcggtgtaagtcgttcgctccaagcttggcgttgtgtagcaaacccccgttccagccga 1594
Db 2490 TTGCGTGTAGGTGCTGTGCTCCAAAGTGGGCTGTGTGACAGAACCCCGCTTCAGCCGA 2549
QY 1595 ccgctgcgcttaccggttaactcgttctgagtcgaaccggtaagaacagactatc 1654
Db 2550 CCGCTGCGCCCTTATCCGCTTACTATCGTCTGAGTCCAAACCGGTAAGACACGACTTATC 2609
QY 1655 gcaactgacagcaacacctgttlaacagagataagcagagcgaggtatgtagcgtgtcac 1714
|||||
```

Accession	Sequence	Position
Dp 2610	GCCACTGGCAGCAGCCACTGTGTAACAGGATTTAGCAADAGGAGGTATGTAGGGGGTGTAC	2668
Qy 1715	agagttcttaagtgtgtgagctactaagctacggtctacactagaagaagatattgtgatatctt	1774
Dp 2670	AGATTCTTTAAGTGGGGGCTACTACGGCTACACTTAAGAGAACAGTATTTGGTATGTG	2729
Qy 1775	cgctctgtctgaagccagttactccttcgaaaaaagtttgttagctctltgatccggtcaaca	1834
Dp 2730	CGCTCTCTGAAGCAGGATTTACCTTCGAAAAAGATGGTGTACTCTCTGTATCGGCAACA	2789
Qy 1835	aaccacgcgcgtgtagagcgtgtgttttttttttgcgaagcgcgagatctaocgcagaataaa	1894
Dp 2790	AACCAACCGCTGGTAGGGGTGTTTGTGTTTGCAAGCAGCATATTCGCGAGAAAAA	2849
Qy 1895	agagtcctcaagaagaatcctcttgatccttctcagcgggtcttgacgtctcagttgaaacgaaa	1954
Dp 2850	AGGATCTCAAGAAAGATCCTTTGATCTTTCTACAGGGGTCTGACGCTCAGTGGAAACGAAA	2909
Qy 1955	ctcacgttaaggatatttgctacttgagcggatatacatatttgaaatgattttga	2007
Dp 2910	CTTCACGTTAAGGGATTTTGGTCTCATGTGATTTATCAAAAAGATCTTCACCTGAA	2962

	RESULT	10		
	AX060344			
	LOCUS			
	AX060344	4597 bp	DNA	PAT
	DEFINITION	Sequence 3 from Patent WO0078358.		22-JAN-2001
	ACCESSION	AX060344		
	VERSION	AX060344.1	GI:12405832	
	KEYWORDS			
	SOURCE			
	ORGANISM	synthetic construct.		
		artificial sequence.		
	REFERENCE	1 (bases 1 to 4597)		
	AUTHORS	Chen,W.		
	TITLE	Hyaluronic acid microspheres for sustained gene transfer		
	JOURNAL	Patent: WO 0078358-A 3 & 28-DEC-2000;		
		The Collaborative Group, Ltd. (US)		
	FEATURES			
	source	Location/Qualifiers		
		1..4597		
		/organism="synthetic construct"		
		/db_xref="taxon:32630"		
		/note="pcDNA3.1/GS vector by Invitrogen Corporation"		
BASE COUNT	1062 a	1214 c	1206 g	1115 t
ORIGIN				

Query Match	12.2%;	Score 988;	DB 9;	Length 4597;
Best Local Similarity	100.0%;	Pred. No. 2.3e-225;		
Matches 988;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

Accession	Sequence	Position
QY 1076	ccacgtcatctaaatgaatcgcccaacgcgcgggaagaacggtttgcgcattgagcgctc	11355
Db 3610	CCACGCTCATTTAATGAATCGGCCCAACGCCGGGGAAGCGGTTTCCGATTGGGCGCTC	36685
QY 1136	ttccgccttcctcgctcaactgactcgctcgctcgcttcgttcgttcgctgacgcgcgtatc	11955
Db 3670	TTCCGCTTCTTCGCTCGCTACTGACTCGCTGCGCTGGCTGCTTTCCGGCTGCGCGACGCGATTAC	37295
QY 1196	aggtcatctcaaaagcggtaatacgtttatcccaagataaaggagataacgcgaagaaagaa	12355
Db 3730	AGGTCACACTCAAAAGCGGCTAATACGGTTATCCACAGATACAGGGGATACAGCGAGAAATAA	37899
QY 1256	catctgacgcaaaagccagcaaaaagccagaaacccctaaaaagccgcgcttctgcgcgt	13155
Db 3790	CATCTGAGCAAAAAGCCAGCAAAAAGGCCAGAAACCTTAAAAAGCCGCGCTTCTGCGCTT	38499
QY 1316	tttccatagcgtccgcgccctctgacgagatcaaaaaatcgacgctcaagtcagagtg	13755
Db 3850	TTTTCATAGGCTCTCGGCCCCCTTACGACGATCCAAATAATCAGCGCTCAAGTCAAGAGGTG	39099
QY 1376	ggcgaacccgcacagcgtataaagataccaagcgtttcccccctggaagctccctctgtgc	14355

Db	3910	GCGAACCCGACAGAGACTATAAAGATACCAAGGGGTTTCCCTCGGAACCTCCTGTGCG	3989
Qy	1436	ctctccctggtccagccctgacggttaccggtataccctgtcgcgctttcccttcgggaag	1485
Db	3970	CTCTCCCTGTTTCCAGACCCCTGCGGCTTACCGGATACCTGTCCGCTTTCCTCGGGAG	4029
Qy	1496	cgtgcgcttcttcaatgcaccagcgtgtagtatactcaattcgcgtgtagtcgttcgctc	1555
Db	4030	CGTGGCGCTTTCATAGCTCACAGCGTGTAGTATGTCAATTGGGTGTAAGTGTTCGCTC	4089
Qy	1556	caagctggtcgtgtgtcagaaccccccggttaagccgacgcgtgcgtccattccggtaa	1615
Db	4090	CAACCTGGCGTGTGTGACCAACCCCCGCTTACGCCCGCACGCTCGCTTATCGGTAA	4149
Qy	1616	ctatcgtccttgatgtcaaacccggttaagaacacttatcgcaactggtgagacgaacatg	1675
Db	4150	CTATCGCTTGTGATGCCAACCCGGTAAACACGACTTATGCCACTGTGACGACGACATTGG	4209
Qy	1676	taacaagatttagaagaagcgaagcgtatgtaaagcgtgtctacagaagtcttctgaagtgtgtcc	1735
Db	4210	TAAACAGATTAGCAGAGCGAGGTATGTAGCGGGTGTACAGAGTTCTTGAAGTGTGTGCC	4269
Qy	1736	taactacggtactcaactagaagaagcaatttgttatctgtcgtctcgtcgtgaagccaattac	1785
Db	4270	TAACTACGGGTACACTTGAAGGAGCACTATTTGGTATCTCGCTGTGTAAGCCAGTTAC	4329
Qy	1796	cttcggaagaagatgtgtatgcctctgtatccgcgaacaacaacacgcgtgtgtagcgtgtg	1855
Db	4330	CTTCGGAAAAAGATTGTGTAGCTCTTGTATTCGCGCAACAAACACCGCGGTAGGGGTGG	4389
Qy	1856	tcttttctgttgcgaagcagcagattacgcgcgcagaaaaaagaatctcaagaagatcctt	1915
Db	4390	TTTTTTTGTGTGCAAGCAGCAGATTTACGGCGAGAAAAAAGATCTCAAGAAAGATCTTT	4449
Qy	1916	gattctttctaggggtgtctgacgtctaagtggaagcaaaaactaagcttaaggaatttgt	1975
Db	4450	GATCTTTTCTACGGGGTCTACCCCTAGTGAACGAAACTACGTTAAGGATTTTGT	4509
Qy	1976	catgaagcgatatactatttaacgttatctagaanaaaataacaacatagtggttccgcgcac	2035
Db	4510	CATGAGCGGATACATTATTGAATGTATTTAAGAAAAATAACAATAATAGGGGTTCCGCGAC	4569
Qy	2036	attctcccgaaaagtgtccactgtagctc 2063	
Db	4570	ATTTCCTCCGAAAAAGTGTCCACTCTACCTGC 4597	

RESULT	11	PAT	07-OCT-1997
LOCUS	163120	4283 bp	DNA
DEFINITION	Sequence 3 from patent US 5661132.		
ACCESSION	163120		
VERSION	163120.1	GI:2480828	
KEYWORDS	.		
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	Unclassified.		
APPHONS	1 (bases 1 to 4283)		
TITLE	Eriksson,E., Andree,C., Swain,W.F. and Macklin,M.D.		
JOURNAL	Wound healing		
FEATURES	Patent: US 5661132-A 3 26-AUG-1997;		
source	Location/Qualifiers		
	1..4283		
	/organism="unknown"		
BASE COUNT	1023 a 1095 c 1114 g 1051 t		
ORIGIN			

Query Match	12.0%;	Score 973;	DB 10;	Length 4283;
Best Local Similarity	83.2%;	Pred. No. 9,1e-22;		
Matches 1216;	Conservative	0;	Mismatches 30;	Indels 215;
				Gaps 1;

D	1517	AGCTGGGGCTCGAGCATGCAAGCTTGAGTATTCCTATAGTGTACCTTAATAATAGCTTGCGCT	1636
Q	1075	-----	1074
D	1637	AATCATGTCATAGCTGTGTTTCGTGTGAAATTTGTTATCCGTCACAAATTCACACAAACA	1696
Q	1075	-----	1074
D	1697	TACGAGCCGGAGACATATAAGTGTAAAGCCTGGGGTGCCCTAATGATGACTAATCACAT	1756
Q	1075	-----	1086
D	1757	TAAATGGGTTGGCTTACTCTCCCGCTTTCCAGTCGGGAAACTGTCCGTGCGCAGCTGCATT	1816
Q	1087	aatgaatcgcgccaacgcgcgagggaagagcggtttgcatlttggcgcttctccgcttcc	1146
D	1817	AATGAATCGGCAACGGCGGGGAGAGCGGGTTGGGTATTTGGGGCGCTTTCGGCTTCT	1876
Q	1147	cgctcaatgactgcgtcgcgctggttccgctgtcggtgagcggtatcaagctcaactaa	1206
D	1877	CGCTCATGTACTCGCTCGCTCGCTGTTCGGCTGGGAGACGGGTATCAGCTCACTCAA	1936
Q	1207	aggcggtaatcaggttatccacagaatcaggggataaagcaggaagaacatgtatgacaa	1266
D	1937	AGCGGGTAAATACGGTTATCCACAGAAATCAGGGGATTAACGAGAAAGAAATGTATAGCAA	1996
Q	1267	aagcgacgcaaaagccaggaacacgtataaaagcgcggttctgtgcglttltccatagc	1326
D	1997	AAGCGCAGCAAAAGCGCCAGGAAACGTTAAAGCCGCTGTCTGGCGTTTTCATAGGC	2056
Q	1327	tcgcgcccccttgacgaagcatcacaaaaatcgaagctcgaagtcgaagtatgagtgagcaaaccca	1386
D	2057	TCCCCCCCCCTGACGAGCATACAAAAATGAGCTCAAGTAAAGGTGGGAAACCCGA	2116
Q	1387	caggactataaagatacagagcglttcccccttggaaagctccccgttgcgtccctcgttc	1446
D	2117	CAGACTATATAAGATACCAAGCGCTTTCCCCGTGAAGCTCCCTCGNGCGCTCCTGTTC	2176
Q	1447	cgacccctgcgcttaccggtatccgttccgcttctcccttcgggaagcgtygcgctt	1506
D	2177	CGACCCCTGCGCTTACCGGATACCTCTTCGCCCTTCTCCCTTGGGAAGCGTGGCGCTTT	2236
Q	1507	ctcaagtctcaagctgtatgatatctcaagttcgggtatgagtgcttcgcgtccaaagctgggct	1566
D	2237	CTCATACTCAGCGTGTAGGTATCTCAGTTCCGTTAGGTCTTTCGCTCCAAAGCTGGGCT	2286
Q	1567	gtgtgcagaaacccccgltcaagccgacgcgtgccttalcggttaactacgtactglttg	1626
D	2297	GTTGTGCAGAAACCCCGCTTACGCCGACCGCGCTGCGCTTATCCGGTAACTATCGCTTGG	2356
Q	1627	agttcaaccccggttaagaacaagacttatgcacacttgcagcagcaacacactgyltaacagata	1686
D	2357	AGTCCAAACCCGTTAAGCAACAGACTTATTCGCCATGTGGACACACCACTGGTAAACAGATTA	2416
Q	1687	gcgaagcagagtatgtatgagcggtgcgtcacaagagttcttgaagtgtgtgcccactacgctt	1746
D	2417	GCAAGGAGAGTATGTAGGGGGGTATCAAGAGTCTTTGAAGTGTGGGCTTAATCTAGGCT	2476
Q	1747	acactagaagaaagfatatttgatctgcgtctctgtcgaagccagtaactctcgaaaaa	1806
D	2477	ACACTAAGAGACAGTATTTGGTATCTGGCGCTCTGCTGAAGCAAGTATACCTTCGGAATAA	2536
Q	1807	gaagtgttagctctcttgaatcggaaaaaacaacacgcgtgtgtagcggtgtgtttttgtt	1866
D	2537	GAGTTGTAGCTCTTGATTCGGCAACAAACCAACGCTGTATGCGGTGGTTTTTTTGT	2596
Q	1867	gcaagcagcagattatcgcgcgagaaaaaagatctcaagaagatccttgaatcttltcta	1926
D	2597	GCAAGCAGCAATTAAGCGCGAGAAAAAAGAGATCTAATAAATATCTTTGATCTTTCTTA	2656
Q	1927	cgaggctcgaagctcaagtgaagcaaaactcaagtttaaggatttggatcatgagcgaat	1986

QY	1987	acataatgataattagata	2007	
Db	2657	CGGGCTCTGACGGCTCAGTGGCAACGAAACTACAGTTAAGGATTTTGCTATGAGATTAT	2716	
QY	1987	acataatgataattagata	2007	
Db	2717	CAAAAGGATCTTCTACCTAGA	2737	
RESULT	13			
LOCUS	AF053407	4050 bp	DNA	circular SYN
DEFINITION	AF053407	Expression vector pMAS-S, complete sequence.		17-OCT-1998
ACCESSION	AF053407			
VERSION	AF053407.1	GI:3746483		
KEYWORDS				
SOURCE				
ORGANISM		Expression vector pMAS-S.		
REFERENCE		Expression vector pMAS-S.		
AUTHORS		artificial sequence; vectors.		
TITLE		1 (bases 1 to 4050)		
JOURNAL		Krieg, A.M., Wu, T., Weeratna, R., Efler, S.M., Love-Homan, L., Yang, L.,		
MEDLINE		Yl, A.R., Short, D., and Davis, H.L.		
REFERENCE		Sequence motifs in adenoviral DNA block immune activation by		
AUTHORS		stimulatory CpG motifs		
JOURNAL		Proc. Natl. Acad. Sci. U.S.A. 95 (21), 12631-12636 (1998)		
FEATURES		98445422		
SOURCE		2 (bases 1 to 4050)		
gene		Wu, T., Efler, S.M., Davis, H.L., Krieg, A.M. and Schorr, J.		
CDS		Direct Submission		
		Submitted (12-MAR-1998) HSD, Loeb Health Research Institute, 725		
		Parkdale Ave., Ottawa, ON K1V 4E9, Canada		
		Location/Qualifiers		
		1..4050		
		/organism="Expression vector pMAS-S"		
		/db_xref="taxon:82305"		
		3351..4031		
		/gene="S"		
		3351..4031		
		/gene="S"		
		/note="derived from Hepatitis B virus surface antigen"		
		/codon_start=1		
		/transl_table=11		
		/product="surface antigen subtype ayw"		
		/protein_id="AAC64340.1"		
		/db_xref="GI:3746484"		
		/translation="MENITSGFLPPLVLVDAGFFLLRLITITQSLDSMTSLNPLAG		
		/TTCVGLGNSQSPSTNSHSPTSCPPCPDGYRWMLRFLIFELILLCLIFLLVLVDG		
		MLPCLPISPSSTSGPCRCMTAGTSMYSCCTKPSDNCNICLIPSSWAFGK		
		FLMMAASARFMSMLSLVPEVFWFGLSPVWLIVIMMWYMGSLVSIISPFLPLPI		
		FFCLMWYI"		
BASE COUNT	993 a	1024 c	949 g	1084 t
ORIGIN				
Query Match	11.9%	Score 968.6;	DB 56;	Length 4050;
Best Local Similarity	85.9%	Pred. No. 1e-220;		
Matches 1155;	Conservative	0;	Mismatches 34;	Indels 156;
				Gaps 1;
QY	736	tgtcactcaaatgctagagctcgctgcatcagcagcctcgagctgcttccttagtgcacgca	855	
Db	40	TCTCGAGCTGACTAGAGCTAGCGTACGATCAACCTCGACTGTCCTTAGTTCACAGCA	99	
QY	856	tctgttgtttgcccctcccccgtgctcttccttagcactgtagagtgacacccacatgctc	915	
Db	100	TCTGTGTTTGGCCCCCTCCCGTGCCTTCTTGACCCCTGGAAAGGTGCACATCCACATGTC	159	
QY	916	ctttcctataaataagagaatgcatcgatgctgctgagtaagtgatcattctatctg	975	
Db	160	CTTTCCTATATAAAGAGAAATTCATCGCATTTGTCGATGAGTGAAGTGCATTCATTCG	219	
QY	976	ggggggtgggggtggggcagacagcaaggggggaggttggaagaacaatagcagcatgct	1035	
Db	220	GGGGGTGGGGGTGGGGCAGACAGCAAGGGGAGGATTTGGAAGCAATAGCAGCATGCT	279	
QY	1036	ggggggtgggggtggggcctatgagc-----	1058	

```
Db 280 GGGGAAGGCGCTCGAGTACGTGGGTAATATGTCATAGCTGTTTCTGTGTGAATTGT 339
1059 -----
Db 340 TATCCGCTCACAATTCACACAAACATACGACCGCGAGACATAAAGTGAAGCCGTGG 399
1059 -----
Db 400 GTGCCCTAATAGTAGTACCTAATCTACATTAATTCGCTTGCGCTCAGTCCGCTTTCCAGT 459
1060 tctgagcgagaaacacccagctgcatlaatgaatcggcaacgcygagagagagcggtc 1119
460 CGGGAACACCTGTCGTGCACCTGACCTCATTAATGAATCGGCCAAGCGGGAGAGCGCGTT 519
1120 tgcgattggcgctcttcgcgtctcctgcgtcaactgactgctgctgctgctgctgcgc 1179
520 TCGGATTTGGGCGCTCTTCGCTTCCCTCGCTCACTGACTGCTGCGCTCGCTGCTGCTGCG 579
1180 tggcgagcgagctatagctcactcaaaagcgtaatacggtatccagaagatcagggg 1239
580 TCGGGGAGGGGATACAGCTCAGCTCAAAAGCGGTAATACGTTATCCACAGAAATCAGGGG 639
1240 ataacgcaagaaagacatgtagcaaaagcgcaagcaaaagcgcaagcaacgttaaaag 1299
640 ATAAACGAGAAAGAACATGTGAGCAAAAGCCAGCAAAAGCCAGAACCGTAAAAAGG 699
1300 cgcgctgctgctgttttccatagagctccgccccctgacgagcatcacaaaaatcgac 1359
700 CCGGTTGGCTGGGTTTTCATATAGGCTCGCCCTCGAGCATCACAAAAATCGAC 759
1360 gctcaagtaagagtgtagcaaaacccgacagagatataaagatccagggcttccctctg 1419
760 GCCTAAGTACAGAGTGGCGCAAAACCGACAGACTATAAATACACAGCGCTTCCCTCG 819
1420 gaagctccctcgctgctctcctgttccgacccctgacgagctacagatattcgcgct 1479
820 GAAGCTCCCTCGGCTCCTCTCTGTTCCGACCTCGCGCTTACCGATCCTGTCCGCT 879
1480 ttctccctcgagaaagcgtagcgcttctcaatgctcaacgctgtagtattcagttcgg 1539
880 TTCTCCCTCGGAGAGCGTGGCGCTTTCATATAGCTCAGCTGTAGTATCTCAGTTCCG 939
1540 tctgagtgctgctcctcaagctgagctgtgtgacgaaaccccgctcagccgacgct 1599
940 TGTAGGCTGTTCCCTCAAGCTGGGCTGTGTGACAGAACCCCGCTTACGCCGACGCT 999
1600 ggcgcttaccggttaactatcgtctgagtaacccgagtaagaacagatattcgcac 1659
1000 GCACCTTATCCGTAATCTGTTGAGTCCAAACCGGTAAGACACAGACTTATTCGCCAC 1059
1660 tggcagcaagcaactgtagcaagagattagcagagcgagtagtagcgctgctacagagt 1719
1060 TGGCAGCAGCACTGGTAAACAGAGATTAGCAGAGAGTATGAGGGGTGCTACAGAGT 1119
1720 tcttgaatgtagtgaactaactacgcttaacagagagagatattgtagtctgcgtc 1779
1120 TCTTGAAGTGTGGCTTAATCTAGGCTACACTGAAGAAAGATTTGGTATCTGCGCTC 1179
1780 tgttgaagcaggttaccttcggaagaaagttgtagctctttagcttcggaagaaaccca 1839
1180 TGTGTGAAGCAGTACCTTCGGAAGAAAGATTGTGACTTTATCTCGGCAAAACAACCA 1239
1840 cgcgctgtagcggtggttttttctgtcgaagcagagattacgcygcaagaaaaaagat 1899
1240 CCGCTGTAGAGCGGTGTTTGTGTTTTCGAAGCAGCAGATTACGCGAGAAAAAAGAT 1299
1900 ctcaagaagatcttctgattcttctcaaggggtctgaagctcagtgtagcaaaactcac 1359
1300 CTCMAAAGATCTTGTGATCTTTTCTACGGGGCTGTGACGCTCAGTGTGAAGAAACATCAC 1359
1960 gttgaagatttgcgtcagtagcg 1984
|||||
```

```
Db 1360 GTTAAGGATTTGTCATGAGCTG 1384
RESULT 14
AF053408
LOCUS AF053408 4227 bp DNA circular SYN 17-OCT-1998
DEFINITION Expression vector pMC616-S, complete sequence.
ACCESSION AF053408
VERSION AF053408.1 GI:3746485
KEYWORDS
SOURCE Expression vector pMC616-S.
ORGANISM Expression vector pMC616-S.
REFERENCE 1 (bases 1 to 4227)
AUTHORS Kriegl,A.M., Wu,T., Weeratna,R., Efler,S.M., Love-Homan,L., Yang,L.,
Yi,A.K., Short,D. and Davis,H.L.
TITLE Sequence motifs in adenoviral DNA block immune activation by
stimulatory CpG motifs
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 95 (21), 12631-12636 (1998)
MEDLINE 98445422
REFERENCE 2 (bases 1 to 4227)
AUTHORS Wu,T., Efler,S.M., Davis,H.L., Kriegl,A.M. and Schorr,J.
TITLES Direct Submission
JOURNAL Submitted (12-MAR-1998) HGD, Loeb Health Research Institute, 725
Parkdale Ave., Ottawa, ON K1V 4E9, Canada
FEATURES
source
1..4227
/organism="Expression vector pMC616-S"
/db_xref="taxon:82306"
3528..4208
/gene="S"
3528..4208
/gene="S"
CDS
/gene="S"
/note="derived from Hepatitis B virus surface antigen"
/codon_start=1
/product="surface antigen subtype ayw"
/protein_id="AAC64341.1"
/db_xref="GI:3746486"
/translation="MENISGFGPLVLVQAGFLLRLITIPQSDNMTSLNFGG
TIVCGONSOPTSNHSPTCPGPGYMMCLRRIFLFLICLLPLVLDYOG
MLPVCPLIGSSSTNSPCRTMTTADGNSMPSCCCTPSGNGCTCIPSPMAFG
FLMEVASRPSMLSLIVFVQNFVGLSPVWLSVIMMYNKPSSLXISPLPLPI
FPCLVNIT"
BASE COUNT 1021 a 1076 c 990 g 1140 t
ORIGIN
Query Match 11.9%; Score 968.6; DB 56; Length 4227;
Best Local Similarity 85.9%; Pred. No. 1e-220;
Matches 1155; Conservative 0; Mismatches 34; Indels 156; Gaps 1;
Db 796 tgcacactaaatgtagagctgctgcatcagctcagctgctgctcctatgtagtccagca 855
40 TCTCGAGTCTAGACTAGACTCCGCTGATCAGCCTCGACTGTGCTTCTTATTCGACCA 99
856 tctgttgcttgcacctcccccgcgtctccttgacctggaaggtgcacctccacgtc 915
100 TCTGTTGTTGGCCCCCTCCCGGCTTCTTGAACCTCGAAGGTGCACTCCACATGTC 159
916 ctctcctaataaattgaggaattgcatgcatgtctgtgtagtgtagtattctattctg 975
160 CTTTCCATTAATTAATGAGGAATTCATCGCATTTGCTGAGTAGGTCTATTCTATTCTG 219
976 gggggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtg 1035
220 GGGGGTGGGTGGGGCGACGACGACGAGGAGGAGATTGGGAAGACATAGCAGCATGCT 279
1036 gggggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtg 1058
280 GGGGAAGGCGCTCGAGTACGTGGGTAATCATGTATAGCTGTTTCTGTGTAATTGT 339
1059 -----
```

```
Dh 340 TATCCGCTCACAAATTCACACACATACGAGCCGGGGAAGCATAAAGTGAAGCCTGGG 399
Qy 1059 -----t 1059
Dh 400 GTGCCTAATGAGTACGTAACCTACATTAAATGCGTTCCTACTGCCCGCTTCCAGT 459
Qy 1060 tctgaagcgaagaagaacagctgcataatgaatcgcgaacgcgcgggaagagcgctt 1119
Dh 460 CGGGAACCTCTCGTCCGACGCTGAATGAATGCGCAACGCCGGGAGAGCGGGT 519
Qy 1120 tgcgtattggcgctcttcgcgtcttcctgcctcgaactgactgcgtgcgtgcgtgc 1179
Dh 520 TGCCTATTGGCGCGCTCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC 579
Qy 1180 tgcggcgagcggttatcaactactcaagaagcggttaactggttatccacagaataagg 1239
Dh 580 TGGCGCGAGCGGTATCAGCTCACTAAAGCGGTAAATACGTTATCCACAGAAATCAGGG 639
Qy 1240 aatacgcagagaagaacaatgtagcaaaagccagcaaaagccagagaacgcgttaaaag 1299
Dh 640 ATAACGCGAGAAAGAACATGTAGAGCAAAAGCCACAAAGGCCAGAACGCTAAAGAG 699
Qy 1300 ccgcgtgcgtgcgttttccataggtccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1359
Dh 700 CCGCGTGTGCTGGCGTTTTCATAGGCTCCGCCCTGACGAGCATCACAATAATCGAC 759
Qy 1360 gctaaagctcaaaagtggagaaaccccgacagaaactataaagacggttttccctctg 1419
Dh 760 GCTCAAGTCAAGAGGTGGGGAACCCGACAGAGACTAAAGATACCAAGCGTTCCTCCCTG 819
Qy 1420 gaagctccctcgtgcgcctcctcgtctcgcgaacctgcgcgttacccgatacctgcgc 1479
Dh 820 GAAGCTCCCTGCTGGCGCTCTCCTGTTCCGACCCCTGCGCTTACCGGATACGTCGCGCT 879
Qy 1480 tctccctcgggaagcgctgcgcgtcttctcaatgctcaacgcgtgtagtatctcagttcg 1539
Dh 880 TTTCCCTCCGGAAGCGCTGGCGCTTCTCATAGCTCAGCGTGTAGGTATCTCACTGCGG 939
Qy 1540 tctaggtcgttcgcgtcccaagctggcgctgtgtagcaaaccccgcttagccgcgcgc 1599
Dh 940 TGTAGGTGTTGCTCCCAAGCTGGGCTGTGTGCACGAACCCCGCTTAGCCCGACGCT 999
Qy 1600 ggcacctaccgtaactactgctctgagtcacaaccgcgtaagaacagactatccgcaac 1659
Dh 1000 GCGCTTATCGGTAACATGTGTGAGTCCACACCGGTAAGACACACTTATGCGCAC 1059
Qy 1660 tggcagcagcaactgtagcaagatagcagagcagaggtatgtaggcggtgctacaggt 1719
Dh 1060 TGGCAGCAGCCACTGGTAAACAGGATTACAGACGAGGATATGAGCGGTGTACAGAGT 1119
Qy 1720 tcttgaagtgtagcctactacaggtactaagaagcagatttgtagtctcgtctc 1779
Dh 1120 TCTTGAATGGTGGCTTAACAGGCTACACTAGAAAGAAATTTGTATCTTCGCTC 1179
Qy 1780 tgcgaagccagttactctgcgaagaagttgtagtcttgcgtccgcgcaacaaca 1839
Dh 1180 TGTGTAACACAGTATACCTTCGGAAGAAAGTTGTAAGCTTGTGATCCGGCAACAAACA 1239
Qy 1840 ccgcgtgtagcggtgttttcttgcgttgcgaagcagcagattacgcgcgaagaaagat 1899
Dh 1240 CCGCTGTAGGCGGTGTTTGTTCGAAGACAGACAGATTACGGCGCAGAAAAAAGAT 1299
Qy 1900 ctcaagaagatccttgcattcttctacaggtctgcagcctcagtgtgaagcaaaataac 1959
Dh 1300 CTCAAGAAGATCTTTGATCTTTCTACGGGGTCTGACGCTAGTGAACGAAACATCAGC 1359
Qy 1960 gttaaagatttgtaataagcgg 1984
Dh 1360 GTTAAAGGATTTGTGTCATGAGCTG 1384
```

```
AF053409
LOCUS AF053409 4625 bp DNA circular SYN 17-OCT-1998
DEFINITION Expression vector pMC650-S, complete sequence.
ACCESSION AF053409
VERSION AF053409.1 GI:3746487
KEYWORDS
SOURCE
ORGANISM Expression vector pMC650-S.
REFERENCE Expression vector pMC650-S.
AUTHORS 1 (bases 1 to 4625).
Krieg,A.M., Wu,T., Weeratna,R., Efler,S.M., Love-Homan,L., Yang,L.,
Yl.A.K., Short,D. and Davis,H.L.
Sequence motifs in adenoviral DNA block immune activation by
stimulatory Cpg motifs
Proc. Natl. Acad. Sci. U.S.A. 95 (21), 12631-12636 (1998)
JOURNAL 98445422
MEDLINE 2 (bases 1 to 4625)
REFERENCE Wu,T., Efler,S.M., Davis,H.L., Krieg,A.M. and Schorr,J.
AUTHORS Direct Submission
JOURNAL Submitted (12-MAR-1998) HCD, Loeb Health Research Institute, 725
Parkdale Ave., Ottawa, ON K1Y 4E3, Canada
FEATURES
source
1..4625
/organism="Expression vector pMC650-S"
/db_xref="taxon:82307"
3926..4606
/gene="S"
3926..4606
/gene="S"
/note="derived from Hepatitis B virus surface antigen"
/codon_start=1
/transl_table=1
/product="surface antigen subtype ayw"
/protein_id="AAC64342.1"
/db_xref="GI:3746488"
/translation="MENVTSQFLGPLVLAQGFLLTRILTIPOSLSQSWTSLNPLG
TTVCLGNSQSPSNNHSPSCPPCPYRWMLRFTIFELLLCLFLVLLDYG
MLPVCPLIPGSSITSPRCMTAGTSMTSCCTKPSDNGCTGIPPSVAFK
FLMEMASAREFMSLSLVPFQWFEVLSPTVWLSVIMMMWGPSTLSLSPFLPLP
FECILMYIT"
BASE COUNT 1084 a 1193 c 1082 g 1266 t
ORIGIN
Query Match 11.9%; Score 968.6; DB 56; Length 4625;
Best Local Similarity 85.9%; Pred. No. 1e-220;
Matches 1155; Conservative 0; Mismatches 34; Indels 156; Gaps 1;
```



```

Db 400 GTGCTAATGAGTGAAGCTAACTCAATTAATTGGTTGCGCTCACTGCCCGCTTTCAGT 459
QY 1060 tctgaagcggaagaaccagctgcatctaataatcgcgaacgcggggaagcggtt 1119
Db 460 CGGGAACCTCTCTCTGCTGCTAATTAATGAATCGCAAGCGCGGGAGAGCGGTT 519
QY 1120 tgcgtatbvgcgtctctccctccctccctccctccctccctccctccctccctcc 1179
Db 520 TGGCTATTGGCGGCTCTTCCGCTTCCGCTCACTGACTGCTGCGCTGCGTGGTGGC 579
QY 1180 tgcgagcgagcggtatcagctcaacgaagcggtatcaggtatccacagaatcaagg 1239
Db 580 TCGCGCGAGCGGTATCAGCTCACTCAAAAGCGGTAATACGGTATCCACAGAAATCAGGG 639
QY 1240 ataacgcaggaagaacaatctgtagcaaaaagccagcaaaaagccaggaaccgtataaag 1299
Db 640 ATAAAGCAGGAAAGAAACATGTGAGCAAAAGGCCAAGCAAGCAAGCAAGCAAGCAAG 699
QY 1300 ccggttgcgtgcttcttccatagctccgccccctgaagcatcacaaaatcgac 1359
Db 700 CCGGCTTGCCTGCTTTTTCATAGGCTCCGCCCTGACGAGCATCACAAAATCGAC 759
QY 1360 gctcaagtcagaagtgagcgaaccggaagactataaagaatacagcggttcccccgt 1419
Db 760 GCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACAGGCGTTTCCCGCTG 819
QY 1420 gaagctccctcgctgcgtctcctcgttccgacctgcggttaccgatacctgtccgct 1479
Db 820 GAAGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 879
QY 1480 ttccctccctccggaagcggtgcgtcttctcaatgctcaagctgtagtactcagttcgg 1539
Db 880 TTCTCCCTTCCGGAAGCGTGGCGCTTCTCATACGCTGCTGAGGATCTCAATTGCG 939
QY 1540 tgaagtcgttgcgttccaaagctggtgtgcaagaaccccggttcaagccgaagcgt 1599
Db 940 TGTAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 999
QY 1600 ggcgcttaccggttaactatcgttctgaagtcgaacccggttaagaacgaactatcgccac 1659
Db 1000 GCGCTTATCCGGTAACCTATCTGATGATCAACCCGGTAAGACAGACTTATCGCCAC 1059
QY 1660 tggagcagccactggtgaacaggaattagcagaaggaattagtagcggtgtctacagat 1719
Db 1060 TGGCAGCAGCCACTGCTAAGCAAGATTAGCAAGCGAGGATGTAGCGGTGCTACAGAGT 1119
QY 1720 tcttgaagtggtgacctaaactacagctacactagaagaacagtaattgtatctgcgtc 1779
Db 1120 TCTTGAAGTGGTGGCTTAACACTAGCGCTAACACTAGAAGACGATTTGGTATCTGCGCTC 1179
QY 1780 tgcgtgaagcagttacactccggaagaagagttgtagctcttgatccggcaacaacca 1839
Db 1180 TGCCTGAAGCCAGTACCTTCGGAAGAGAGTGGTACCTCTTGTATCCGGCAACCAACA 1239
QY 1840 ccggttgaagcaggtggtttttttgttgcgaagcagagattacgcggaagaaaaagat 1899
Db 1240 CCGCTGCTAGGAGGAGTGTGTTTTTTTGTGCAAGCAGCAGATTACGCGCAAGAAAAAGAT 1299
QY 1900 ctcaagaagatccttltgatcttltctaaggggtcgaagctcaagtggaacgaacactaac 1959
Db 1300 CTCAAGAAGATCCTTTTGTATCTTTCTACGGGGGTCTGACGCTCACTGGAACGAAAACTCAC 1359
QY 1960 gtttaaggaatttggtcatagagcg 1984
Db 1360 GTTAAAGGATTTTGGTCATGAGCTG 1384

```

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 3, 2001, 16:01:04 ; Search time 444.17 Seconds

(without alignments)
11471.782 Million cell updates/sec

Title: US-09-778-516-1

Perfect score: 8115

Sequence: 1 gatctacggcgccagatatcac.....gcgtctcgtacgacagtc 8115

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 730101 seqs, 313950809 residues

Total number of hits satisfying chosen parameters: 1460202

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_0601:*

- 1: /SIDSL/gcgdata/geneseq/geneseqn/NA1980.DAT:*
- 2: /SIDSL/gcgdata/geneseq/geneseqn/NA1981.DAT:*
- 3: /SIDSL/gcgdata/geneseq/geneseqn/NA1982.DAT:*
- 4: /SIDSL/gcgdata/geneseq/geneseqn/NA1983.DAT:*
- 5: /SIDSL/gcgdata/geneseq/geneseqn/NA1984.DAT:*
- 6: /SIDSL/gcgdata/geneseq/geneseqn/NA1985.DAT:*
- 7: /SIDSL/gcgdata/geneseq/geneseqn/NA1986.DAT:*
- 8: /SIDSL/gcgdata/geneseq/geneseqn/NA1987.DAT:*
- 9: /SIDSL/gcgdata/geneseq/geneseqn/NA1988.DAT:*
- 10: /SIDSL/gcgdata/geneseq/geneseqn/NA1989.DAT:*
- 11: /SIDSL/gcgdata/geneseq/geneseqn/NA1990.DAT:*
- 12: /SIDSL/gcgdata/geneseq/geneseqn/NA1991.DAT:*
- 13: /SIDSL/gcgdata/geneseq/geneseqn/NA1992.DAT:*
- 14: /SIDSL/gcgdata/geneseq/geneseqn/NA1993.DAT:*
- 15: /SIDSL/gcgdata/geneseq/geneseqn/NA1994.DAT:*
- 16: /SIDSL/gcgdata/geneseq/geneseqn/NA1995.DAT:*
- 17: /SIDSL/gcgdata/geneseq/geneseqn/NA1996.DAT:*
- 18: /SIDSL/gcgdata/geneseq/geneseqn/NA1997.DAT:*
- 19: /SIDSL/gcgdata/geneseq/geneseqn/NA1998.DAT:*
- 20: /SIDSL/gcgdata/geneseq/geneseqn/NA1999.DAT:*
- 21: /SIDSL/gcgdata/geneseq/geneseqn/NA2000.DAT:*
- 22: /SIDSL/gcgdata/geneseq/geneseqn/NA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3497.2	43.1	5059	21	AA236227
2	3110.6	38.3	3203	11	AA006760
3	1245.2	15.3	3853	19	AAV40006
4	1152.4	14.2	5306	21	AAA53878
5	1148.2	14.1	14455	19	AAV32374
6	1148.2	14.1	14455	21	AAA59050
7	1140	14.0	5377	21	AAA53872
8	1131.4	13.9	4645	21	AAA53870
9	1082	13.3	5446	19	AAV38297
10	1071.6	13.2	5245	21	AAA53871
11	1070.2	13.2	4026	19	AAV40007

12	1068	13.2	8705	20	AA237778	Vector pShuttle DN
13	1058.4	13.0	6050	21	AAA47797	PMR/SBO-S vaccine
14	1058.4	13.0	8002	21	AAA47796	PMR/SBO-M vaccine
15	1036.2	12.8	3796	21	AAA27831	Vector plasmid PCM
16	988	12.2	4597	22	AA24901	Nucleotide sequenc
17	984.6	12.1	8136	18	AA263236	Plasmid PCR(TM)31a
18	981.8	12.1	6253	20	AA208454	AAV vector sequenc
19	981.8	12.1	6280	20	AA208455	AAV vector sequenc
20	981.8	12.1	6280	20	AA208456	AAV vector sequenc
21	981.8	12.1	6365	21	AA245925	Nucleotide sequenc
22	981.8	12.1	6981	21	AA245931	Nucleotide sequenc
23	981.8	12.1	7054	21	AA245927	Nucleotide sequenc
24	981.8	12.1	7405	21	AA245926	Nucleotide sequenc
25	973	12.0	4283	18	AA286449	DNA encoding hgh
26	973	12.0	4283	19	AA286610	Epidermal growth f
27	973	12.0	4283	21	AA263237	Plasmid PMR61630 n
28	961.4	11.8	6200	21	AA249493	Human alpha-L-idur
29	959.6	11.8	6115	18	AA263235	Eukaryotic express
30	948.4	11.7	5857	21	AA295423	Expression vector
31	935.6	11.5	5432	21	AA289476	Transgenic APP DNA
32	932.8	11.5	5732	21	AA271429	Expression vector
33	932	11.5	5729	21	AA271430	Expression vector
34	921.8	11.4	6281	22	AA284206	Plasmid Glambda-1A
35	921.8	11.4	6284	19	AA241427	Plasmid Hu19AHcpd
36	921.8	11.4	6285	17	AA234109	Anti-IL-5 humanise
37	921.8	11.4	6285	19	AA203503	Plasmid pCD115H2HC
38	912.6	11.2	6094	21	AA271428	Expression vector
39	908.2	11.2	5957	21	AA255464	Destination vector
40	908.2	11.2	5957	21	AA255467	Destination vector
41	906.4	11.2	2704	20	AA206758	Vector pUC28 nucle
42	906.4	11.2	3130	19	AA246334	Universal reporter
43	906.4	11.2	3364	21	AA239495	Transgenic unc-119
44	906.4	11.2	3789	21	AA239782	Plasmid pBIMMN nu
45	906.4	11.2	4173	19	AA232663	Plasmid pGP-TT re

ALIGNMENTS

RESULT 1	
ID AA236227	standard: DNA; 5059 BP.
XX	
AC AA236227;	
XX	
DT 22-FEB-2000	(first entry)
XX	
DE	DNA encoding a permease an beta-galactosidase.
XX	
KW	Lactose operon; permease; beta-galactosidase; mutant; yoghurt;
KW	beta-galactosidase activity; lactose fermentation;
KW	fermented milk product; Streptococcus thermophilus; ds.
OS	
XX	Lactobacillus bulgaricus.
XX	
FT	
FT	Key
FT	Location/Qualifiers
FT	122..1873
FT	/tag= a
FT	/product= "permease"
FT	/note= "encodes AA236220"
FT	1877..4519
FT	/tag= D
FT	/product= "beta-galactosidase"
FT	/note= "encodes AA236221"
XX	
FR2778921-AL	
XX	
PN	26-NOV-1999.
XX	
PD	
XX	
PF	22-MAY-1998;
XX	98PR-0006456.
XX	
PR	22-MAY-1998;
XX	98PR-0006456.

PA (DANO-) CITE DANONE SA GERVAIS.
 XX
 PI Benbadis L, Brignon P, Gendre F;
 XX
 DR WPI: 2000-026146/03.
 DR P-PSDB: AMV53620, AMV53621.
 XX
 PT New mutant *Lactobacillus bulgaricus* strain lacking beta-galactosidase
 PT activity useful for making fermented milk products, especially yoghurt
 XX
 PS
 PS Disclosure: Page 15-22; 32pp; French.
 XX
 CC The present sequence represents a lactose operon of *Lactobacillus*
 CC *bulgaricus*, and encodes a permease protein and beta-galactosidase
 CC protein. The specification describes mutant *L. bulgaricus* strain lacking
 CC beta-galactosidase activity, because it has a termination (nonsense)
 CC mutation in at least one of the coding sequences of the lactose operon.
 CC The resulting mutant strain is unable to ferment lactose. Since the
 CC mutant strain cannot ferment lactose, fermented milk products produced
 CC using it do not have to be cooled to prevent post-acidification. The
 CC mutant *L. bulgaricus* strain of the invention is used for producing a
 CC fermented milk product, especially yoghurt. Especially, milk is
 CC fermented with a culture of the mutant strain and optionally
 CC *Streptococcus thermophilus* in the presence of at least one sugar
 CC assimilable by the mutant strain, preferably glucose.
 XX
 Q0 Sequence 5059 BP; 1229 A; 1341 C; 1323 G; 1166 T; 0 other;

Query Match	43.18;	Score 3497.2;	DB 21;	Length 5059;
Best Local Similarity	99.58;	Pred. No. 0;		
Matches 3518;	Conservative 0;	Mismatches 13;	Indels 3;	Gaps 1;

QY	4584	ccgtataaaagttagcgagcagctcttgccttcgctgcgcagcagcgtatagttctacgcgc	4643
Db	1479	ccgcacaaaaagttatagtcgcagcagtgctttgacctgtgtccacagacgtatagttctacgcgc	1538
QY	4644	catttcgcggtactgtccgcgcactctgcgcacaagaccgcgaactcgatctccttcgtaaatg	4703
Db	1539	catttcgcggtactgtccgcgcactctgcgcacaagaccgcgaactcgatctccttcgtaaatg	1588
QY	4704	aaacatggggtctgtgctcttctatccactctgcgcctgcgcacggttcaaatlaaacggtgactg	4763
Db	1599	aaacatggggtctgtgctcttctatccactctgcgcctgcgcacggttcaaatlaaacggtgactg	1658
QY	4764	gctttctgcagctatgtctgaagagggcagccaggtatgaagccggtccagcagatctctggaat	4823
Db	1659	gctttctgcagctatgtctgaagagggcagccaggtatgaagccggtccagcagatctctggaat	1718
QY	4824	tcctgcgacccgcgcgtatcaaacagccaaagctctgcagcagccggttaatctgtacgcttcatca	4883
Db	1719	tcctgcgacccgcgcgtatcaaacagccaaagctctgcagcagccggttaatctgtacgcttcatca	1778
QY	4884	acagcgcgaacattcttcacaaatagcagcagatctctcttcgcgaatcgcgcacagcgttccaaagcc	4943
Db	1779	acagcgcgaacattcttcgcgaatagcagcagatctctcttcgcgaatcgcgcacagcgttccaaagcc	1838
QY	4944	tgtatgtatgtatctccaagtttagaagggaagaaatagaataatgacataatgtatgaataag	5003
Db	1839	tgtatgtatgtatctccaagtttagaagggaagaaatagaataatgacataatgtatgaataag	1898
QY	5004	aaaaaaagatttgaccagcgcagactctgcgtctatctgcataccgcgaagaatttaccagaatgca	5063
Db	1899	aaaaaaagatttgaccagcgcagactctgcgtctatctgcataccgcgaagaatttaccagaatgca	1958
QY	5064	ataaactctcccccgcacactccgcacatactgactctcttccaaaagccaggaagaagaacttggagaagg	5123
Db	1959	ataaactctcccccgcacactccgcacatactgactctcttccaaaagccaggaagaagaacttggagaagg	2018
QY	5124	gcaagatccagatttagtgcagctccctctgcagcgggagactgctgatactacgtacgtgaaaagc	5183
Db	2019	gcaagatccagatttagtgcagctccctctgcagcgggagactgctgatactacgtacgtgaaaagc	2078

QY	5184	gcccagggacagctaaccttctatgcccagaaagactcttgacatgacaaattttaagtatgca	5243
Db	2079	gcccagggacccagctaaccttctatgcccagaaagactcttgacatgacaaattttaagtatgca	2138
QY	5244	aagtaacccggacaaccttgaaactgcaagagctcttgcccagcccagatgctgcaagctcat	5303
Db	2139	aagtaacccggacaaccttgaaactgcaagagctcttgcccagcccagatgctgcaagctcat	2188
QY	5304	atccatgagcagcgagctgaaagagatcttcgccgcccaatctccaaagcaaaaatccgcctcg	5363
Db	2199	atccatgagcagcgagctgaaagagatcttcgccgcccaatctccaaagcaaaaatccgcctcg	2258
QY	5364	ctctctatgctcaagtaattcttgaaactctgaaagactcttcgggacaagaagaaacagcttga	5423
Db	2259	ctctctatgctcaagtaattcttgaaactctgaaagactcttcgggacaagaagaaacagcttga	2318
QY	5424	agttgacggggcgagcaaacagccatctatgctcgtgacaagcgccactctgcgcctacg	5483
Db	2319	agttgacggggcgagcaaacagccatctatgctcgtgacaagcgccactctgcgcctacg	2378
QY	5484	ggggaagactccttaccaccaagcgagttaatgcttacccaagtctctcagaagaagaata	5543
Db	2379	ggggaagactccttaccaccaagcgagttaatgcttacccaagtctctcagaagaagaata	2438
QY	5544	accgcctgagcgctgagctctctacaagtatcttcgcgcctccgcctgagcgagacagagact	5603
Db	2439	accgcctgagcgagctgagctctctacaagtatcttcgcgcctccgcctgagcgagacagagact	2488
QY	5604	ctctgcgcatgctcgtgttctgctcagatcagtgactcttcagagccaaagccgcgtctgcaact	5663
Db	2499	ctctgcgcatgctcgtgttctgctcagatcagtgactcttcagagccaaagccgcgtctgcaact	2558
QY	5664	tggggagaccttaagaacttaacggcgagactctacccaataacacaaaagaagaagctgagag	5723
Db	2559	tggggagaccttaagaacttaacggcgagactctacccaataacacaaaagaagaagctgagag	2618
QY	5724	tcgaaagccaatattgctcctaacgcctctgcacaatgccaagctcttaagctggaagtgcgggata	5783
Db	2619	tcgaaagccaatattgctcctaacgcctctgcacaatgccaagctcttaagctggaagtgcgggata	2678
QY	5784	gtgaaagtgaactctgtctgcggaaaagctgagcccaatccgaagcgagcagcgtcggaattca	5843
Db	2679	gtgaaagtgaactctgtctgcggaaaagctgagcccaatccgaagcgagcagcgtcggaattca	2738
QY	5844	ctctgagctgaattctgcagatgctgcgcctgcggagcgcggaagaagcctaaccttaccagctcc	5903
Db	2739	ctctgagctgaattctgcagatgctgcgcctgcggagcgcggaagaagcctaaccttaccagctcc	2788
QY	5904	gctcgtatattatcacagcgagcgagccctcttgaagttagccggcgaggaagtgggttcc	5963
Db	2799	gctcgtatattatcacagcgagcgagccctcttgaagttagccggcgaggaagtgggttcc	2858
QY	5964	gcaacatttgaacttaaaagagaggaattatgtactcttaacggcgagcgaggtctgctcttaacag	6023
Db	2859	gcaacatttgaacttaaaagagaggaattatgtactcttaacggcgagcgaggtctgctcttaacag	2918
QY	6024	gggccaacccggcagaagaattctggaacagtaagtctgtctgggctatccacggaaagaggaatga	6083
Db	2919	gggccaacccggcagaagaattctggaacagtaagtctgtctgggctatccacggaaagaggaatga	2978
QY	6084	tcctgggacataagacacataaagcggaaagcaaatcaaatgtcttcgcgtctctaactacc	6143
Db	2979	tcctgggacataagacacataaagcggaaagcaaatcaaatgtcttcgcgtctctaactacc	3038
QY	6144	cgaagccgctccctcttcttacggcgctctgtaacaagaagtaagcctttagctatgtagag	6203
Db	3039	cgaagccgctccctcttcttacggcgctctgtaacaagaagtaagcctttagctatgtagag	3088
QY	6204	ctaacctggaagaccacgagcactctgggaaaaaagtctgggggacagaaagatccatgactca	6263
Db	3099	ctaacctggaagaccacgagcactctgggaaaaaagtctgggggacagaaagatccatgactca	3158

QY 6264 atgtccagggcatgataccagcatgtgctgaggagccgtatacccggtgtgaagaacatga 6323
|||||
Db 3159 atgtccagggcatgataccagcatgtgctgaggagccgtatacccggtgtgaagaacatga 3218
QY 6524 tggctcgggacaaagaacatgcttcaatcctaactgtgctttaggcgaatgagcttaca 6383
|||||
Db 3219 tggctcgggacaaagaacatgcttcaatcctaactgtgctttaggcgaatgagcttaca 3278
QY 6384 ccggacatgtcttctggcccaaatggtctattacgtctccggaagagctgataccgacgggtctc 6443
|||||
Db 3279 ccggacatgtcttctggcccaaatggtctattacgtctccggaagagctgataccgacgggtctc 3338
QY 6444 agccactatgaagggtgtagcccaacacgggaagtttgagacgacccacagattgaagacc 6503
|||||
Db 3339 agccactatgaagggtgtagcccaacacgggaagtttgagacgacccacagattgaagacc 3398
QY 6504 ggaatgtatgtcccgccaaaggtaattgaagaatacttgacaataaaaccaacgaagccat 6553
|||||
Db 3399 ggaatgtatgtcccgccaaaggtaattgaagaatacttgacaataaaaccaacgaagccat 3458
QY 6564 ttactcaagtgtgaatacgtctacgcatggtgcaactccgtctggtgacatggtcgccataca 6623
|||||
Db 3459 ttactcaagtgtgaatacgtctacgcatggtgcaactccgtctggtgacatggtcgccataca 3518
QY 6624 cggccctggaaaaataacccccactacacagggcggtcttcatcttggaactggaattgaacaa 6683
|||||
Db 3519 cggccctggaaaaataacccccactacacagggcggtcttcatcttggaactggaattgaacaa 3578
QY 6684 gactggaaaaaagacgggacacgtgcttataggggcgacatcgatgacggacgcaacagact 6743
|||||
Db 3579 gactggaaaaaagacgggacacgtgcttataggggcgacatcgatgacggacgcaacagact 3638
QY 6744 atgaaattctcggggaagacggcgctgtcttctgctgacccggaactggaatcgccgaactgtgcta 6803
|||||
Db 3639 atgaaattctcggggaagacggcgctgtcttctgctgacccggaactggaatcgccgaactgtgcta 3698
QY 6804 atgcaagggccctttagcccaacacttaagttagaagtaaaagatgggcagatcttccca 6863
|||||
Db 3699 atgcaagggccctttagcccaacacttaagttagaagtaaaagatgggcagatcttccca 3788
QY 6864 aaaaagcaaatltaattacacaacagctcatcttactacttcttgactagttcttggctcg 6923
|||||
Db 3759 aaaaagcaaatltaattacacaacagctcatcttactacttcttgactagttcttggctcg 3818
QY 6924 atggcaagttgacctaccagagccggcgctctgaccttggcctggagagctggcgcaatccg 6993
|||||
Db 3819 atggcaagttgacctaccagagccggcgctctgaccttggcctggagagctggcgcaatccg 3878
QY 6984 ggaaccttgcctgctgcttggccggaatcgctgataaaaaaggggaggtcgcttaccggg 7043
|||||
Db 3879 ggaaccttgcctgctgcttggccggaatcgctgataaaaaaggggaggtcgcttaccggg 3938
QY 7044 taacggcccaacttaaaagaagacttgcttggcggaatgaagggttcaactgtgctgaag 7103
|||||
Db 3939 taacggcccaacttaaaagaagacttgcttggcggaatgaagggttcaactgtgctgaag 3998
QY 7104 cagagaagaagttagctccaaaagctgcccgaatttaagccggaaggcgccagatttagttg 7163
|||||
Db 3999 cagagaagaagttagctccaaaagctgcccgaatttaagccggaaggcgccagatttagttg 4058
QY 7164 attccgaactacaacttagggcctgtaaaaggaataacttccaattctcttcccaagtga 7223
|||||
Db 4059 attccgaactacaacttagggcctgtaaaaggaataacttccaattctcttcccaagtga 4118
QY 7224 agggctggccgggttccctcaagatagccggtaggaatacttgaacgggttcgggaat 7283
|||||
Db 4119 agggctggccgggttccctcaagatagccggtaggaatacttgaacgggttcgggaat 4178
QY 7284 ttaccttctggcggtccctgtagcaacagacggcggaagctggtttagcggtatgactgtg 7343
|||||
Db 4179 ttaccttctggcggtccctgtagcaacagacggcggaagctggtttagcggtatgactgtg 4238
QY 7344 ccgggtgggaataatgcccggcgaagtgcggcttgtaagaacatcagctgtaggttcaagg 7403

Db 4229 cccggtgggaataatgcccggcgaagtatgcccgtcttgaagaacatcagctgtaggttcaagg 4298
|||||
QY 7404 aagactccgttttggctcaagaactcctttagctgtgctgtgcttcaaggtgatttaa 7463
|||||
Db 4299 aagactccgttttggctcaagaactcctttagctgtgctgtgcttcaaggtgatttaa 4358
QY 7464 ccgtgacctatgaagttagatgagcggtggaagattgtgctgtaacggtgacttcccaaggcg 7523
|||||
Db 4359 ctgtgacctatgaagcttagagcggtggaagattgtgctgtaacggtgacttcccaaggcg 4418
QY 7524 cggagaagaactgtgctctcttggcacaagcttggcttgaaccttgcccttgcacaagaactga 7583
|||||
Db 4419 cggagaagaactgtgctctcttggcacaagcttggcttgaaccttgcccttgcacaagaactga 4478
QY 7584 ccgaattacgctactatgctgtgtagcctaatagtagagctacccagacggcttgaaggtta 7643
|||||
Db 4479 ccgaattacgctactatgctgtgtagcctaatagtagagctacccagacggcttgaaggtta 4538
QY 7644 attaccgtggatcttaccagggagcggtaaanaagaacttagcccata---tcgtccgc 7700
|||||
Db 4539 attaccgtggatcttaccagggagcggtaaanaagaacttagcccatactgctgcgc 4598
QY 7701 aggaacggggcaaccggagcgaagttcgctgtgtaccagctcttggatgaagggcggt 7760
|||||
Db 4599 aggaacggggcaaccggagcgaagttcgctgtgtaccagctcttggatgaagggcggt 4658
QY 7761 tggaaattacggccaatggggcagacttgaactgtctgtcttggccaatctgcggccc 7820
|||||
Db 4659 tggaaattacggccaatggggcagacttgaactgtctgtcttggccaatctgcggccc 4718
QY 7821 aaattgaagcgagcgagccacgctttagaactgactataaataacttgggttagagct 7880
|||||
Db 4719 aaattgaagcgagcgagccacgctttagaactgactataaataacttgggttagagct 4778
QY 7881 taagcgcccaatggtgggtcgcggtggaatgaactcctcgtggggcagaaggttccaccgggaat 7940
|||||
Db 4779 taagcgcccaatggtgggtcgcggtggaatgaactcctcgtggggcagaaggttccaccgggaat 4838
QY 7941 tctgcttgaatgtctcaaaaagcccgccagcttgcgcttggatgattcaagcccttactaa 8000
|||||
Db 4839 tctgcttgaatgtctcaaaaagcccgccagcttgcgcttggatgattcaagcccttactaa 4898
QY 8001 aataaagtctacaatttgaacttaacagatgaataatttagtaaaagcagaagttagga 8060
|||||
Db 4899 aataaagtctacaatttgaacttaacagatgaataatttagtaaaagcagaagttagga 4958
QY 8061 agatggcaacgatacagaagaatgccaagggcagccggcggtgctgcttagcgaaggt 8114
|||||
Db 4959 agatggcaacgatacagaagaatgccaagggcagccggcggtgctgcttagcgaaggt 5012

RESULT 2
AAQ06760
ID AAQ06760 standard; DNA: 3203 BP.
XX
AC AAQ06760;
XX
DT 25-FEB-1991 (first entry)
XX
DE Sequence of beta-galactosidase gene.
XX
KM PBR329; pBG1; ds.
XX
OS Lactobacillus bulgaricus SBR0034.
XX
FH Key Location/Qualifiers
FT CDS 285..3083
FT /*tag= a
PN JF02261383-A.
XX
PD 24-OCR-1990.

QY	7796	tcgtcttgcacatattcttcgcccacaaatgtgaagcgcgagaccacgcttttgaactgact	7855
Db	2877	ttctcttgcacatattcttcgcccacaaatgtgaagcgcgagaccacgcttttgaactgact	2936
QY	7856	aacattacacttgggttaagaccttaagcgcgcacagatgggggtctcgccgggagtactcc	7915
Db	2937	aacattacacttgggttaagaccttaagcgcgcacagatgggggtctcgccgggagtactcc	2996
QY	7916	tgggggcagaaggtccaccgcccggaattctgccttgaatgctcacaagaaccgcgcgttcgc	7975
Db	2997	tgggggcagaaggtccaccgcccggaattctgccttgaatgctcacaagaaccgcgcgttcgc	3056
QY	7976	ctgtgtactcaagcccttctaataaataatgctacaatgtgacttaacagagatgaat	8035
Db	3057	ctgtgtactcaagcccttctaataaataatgctacaatgtgacttaacagagatgaat	3116
QY	8036	ttaataaagaagaagcagatgaggaagatggcaagcatagaagaagtccaaagcagcg	8095
Db	3117	ttaataaagaagaagcagatgaggaagatggcaagcatagaagaagtccaaagcagcg	3176
QY	8096	gcgtgtcgtcagcagcgt	8114
Db	3177	gcgtgtcgtcagcagcgt	3195
RESULT 3			
AAV40006	AAV40006 standard; DNA; 3853 BP.		
AC	AAV40006;		
XX	15-FEB-1999 (first entry)		
DE	Plasmid pCTM.		
KW	E2f; transcription factor; human; retinoblastoma protein RB;		
KW	bladder cancer; restenosis; angioplasty; diabetic retinopathy;		
KW	thyroid hyperplasia; Grave's disease; psoriasis;		
KW	benign prostatic hypertrophy; Li-Fraumeni syndrome;		
KW	peripheral vascular disease; therapy; plasmid pCTM; ss.		
XX	Chimeric - cytomegalovirus.		
OS	Chimeric - mastadenovirus.		
OS	Chimeric - bacteriophage T7.		
OS	Chimeric - bacteriophage SP6.		
OS	Chimeric - rhesus macaque polyoma virus.		
XX	Chimeric - Bos taurus.		
XX	Key		
FT	Location/Qualifiers		
FT	promoter		
FT	209..864		
FT	/tag= a		
FT	/note= "CMV promoter"		
FT	907..1131		
FT	misc.feature		
FT	/tag= b		
FT	/function= tripartite leader sequence		
FT	1132..1149		
FT	promoter		
FT	/tag= c		
FT	/note= "SP6 promoter"		
FT	1679..3853		
FT	misc.feature		
FT	/tag= d		
FT	/note= "pUC19 backbone H3 to AattI"		
FT	complement (2857..3717)		
FT	/tag= e		
FT	/note= "AMP-OR"		
PN	W09821228-A1.		
XX	22-MAY-1998.		
XX	13-NOV-1997.		
XX	97MO-US21821.		
XX	14-FEB-1997.		
XX	97OS-0801092.		
PR	15-NOV-1996.		
PR	96US-0751517.		


```

QY 383 tatiatgcatcgtattaccatggtgatgcggttttggcagttacatcattggcgctggat 442
Db 500 tatiatgcatcgtattaccatggtgatgcggttttggcagttacatcattggcgctggat 559
QY 443 agcggttctgactcaagggtattcccaagttcccaaccattgagctcaatgggaattgt 502
Db 560 agcgcttctgactcaagggtattcccaagttcccaaccattgagctcaatgggaattgt 619
QY 503 ttggcaccaaatcaaaagggaacttcccaaatgtcgttaacaaact-----ccggccc 554
Db 620 ttggcaccaaatcaaaagggaacttcccaaatgtcgttaacaaactgtcgtccggccc 679
QY 555 attgacgcaaatgggcggttaaggcgtgtacacggttgggaggttctatatgaacagagctctcg 614
Db 680 gttagcgaacaaatgggcggttaaggcgtgtacacggttgggaggttctatatgaacagagctctcg 739
QY 615 gctactatgagaacccactgcttactatgctatcgaaatlaatacagactcaactatagga 674
Db 740 agtgaacggttagatcatcgtgaattcttgacgaactactgtattaaagcgcatagaggtctcc 799
QY 675 gaccgaagctgtgatacggagctcgatccactagttaaagcgccgacagtggtgtg-aaatt 733
Db 800 tgcagaatcttctatagtgtaacaactatcgatttccacacatatatacgaagccgagtgttaatt 859
QY 734 ctgcaagatcatcacaacacttgcgacgctcgagcatgcatctagagggccctattctat 793
Db 860 gtcaacagctcatgtatgaacgtcccgaggagacagaagccccaacatgctcgtcgtagtaata 919
QY 794 agtctcaactaaatgcttgaagctcgtctatcagctcgtacgtgtcgtctctagtctcagc 853
Db 920 cgaactcatatagggcgagatgtgagttctcgtactcgtacttaagagag----- 965
QY 854 catctgttcttgcgcccccccggtctctccctctgacccctggaaggttgcactccactg 913
Db 966 -----gctatctgcccaggttaccgtctgacgctgctgaagaaagagt 1000
QY 914 tcccttccataaataatgagaacttgcacatgctgtctcgtagtggttcaattctatcc 973
Db 1001 ttaagagagcgccgaacaaagcgtcatgagccgaagtgtgcgagccgactcttcccactcg 1060
QY 974 tggggggtgggtgtgggagacagcaagggtgagagatttgggaagacatatagcagctatg 1033
Db 1061 gtgcatgtcggatataatagtcgcaagcaaccgacacttgcgacggtgagtcgagccacg 1120
QY 1034 ctggggatgcggtgtgcctctatgctctgagggcggaagaaacccgcgtcatatgaat 1093
Db 1121 atggc-----tcggcggttagaagatccacagaagcggtgtgtgtgcacatga 1166
QY 1094 cggccaacgcgcggggagagagcggttgcgtatttggcgctcttccgcttccctgcatac 1153
Db 1167 tgcggttagctatagttggtctcaagtacgaagcgagcagagcttggcgcgcccaagc 1226
QY 1154 tgaactcgtgtgcgtcgtcgtctcgtctgagcgagcggtatcagctactcaaaagcggt 1213
Db 1227 ggtcggagaggtgctcgcg-agaacgggtgtgcataagaaatgcatcaagcatatagcgct 1285
QY 1214 aatcgggttatcccaacagatcaggggataacgcaagcaagaaagacatgttagcaaaagcgca 1273
Db 1286 agatccctgtcagtagtctgagatcgtcagac-----catgttagcaaaagcgca 1334
QY 1274 gcaaaagcgcaagaaacgttaaaagggcgcggttgcgtcgtttttccataggtccgccc 1333
Db 1335 gcaaaagcgcaagaaacgttaaaagggcgcggttgcgtcgtttttccataggtccgccc 1394
QY 1334 cccctgacgagatatacaaaaatcgacgtctcaagtcaaggttggcgaaacccgcagagact 1393
Db 1395 cccctgacgagatatacaaaaatcgacgtctcaagtcaaggttggcgaaacccgcagagact 1454
QY 1394 ataaagataccagcggttcccccgtgaagctccctcgtcgtcctctcgttttccgaacct 1453
Db 1455 ataaagataccagcggttcccccgtgaagctccctcgtcgtcctctcgttttccgaacct 1514
QY 1454 gccgcttaccggtatcgtctcgccttctcccttcggaagcggtgcgcttctcctaatg 1513

```

```

Db 1515 gccgcttaccggtatccctgtccgccttctccctccctcgggaagcggtgcgcttctcctatag 1574
QY 1514 ctcaagcgttagtatctcttaagttcgtgtgtaggtcgttccgtccaagctgtgctgtgtga 1573
Db 1575 ctcaagcgttagtatctcttaagttcgtgtgtaggtcgttccgtccaagctgtgctgtgtga 1634
QY 1574 cgaacccccgcttcagcccgacgctgtgcgcttaccggttaactatcgtctgagttccaa 1633
Db 1635 cgaacccccgcttcagcccgacgctgtgcgcttaccggttaactatcgtctgagttccaa 1694
QY 1634 cccggtlaagacacgaactatcgcacacttgcagcagccacattgtaacagattagcagagc 1693
Db 1695 cccggtlaagacacgaactatcgcacacttgcagcagccacattgtaacagattagcagagc 1754
QY 1694 gaggtatgtgaagcggtgtgtcagcagagttcttgaaggtgtggtgcttaactagcgttaactag 1753
Db 1755 gaggtatgtgaagcggtgtgtcagcagagttcttgaaggtgtggtgcttaactagcgttaactag 1814
QY 1754 aaggaacagttattgttatctgtcgtctgtcgttgaagccagttaccttcggaaaaagattg 1813
Db 1815 aaggaacagttattgttatctgtcgtctgtcgttgaagccagttaccttcggaaaaagattg 1874
QY 1814 tagctcttgcacccggaacaaacacacgcgtgtgtagcgtgtgtttttgttttgcgaagca 1873
Db 1875 tagctcttgcacccggaacaaacacacgcgtgtgtagcgtgtgtttttgttttgcgaagca 1934
QY 1874 ggaatttcgacgcaaaaaaagatctcaagaagatcccttgcattcttctcagggggtc 1933
Db 1935 gcaagttacgcgcaaaaaaagatctcaagaagatcccttgcattcttctcagggggtc 1994
QY 1934 tgacgtcagttggaacggaacacacacgttaagagatttgcgtcatgagcggaatacatatt 1993
Db 1995 tgacgtcagttggaacggaacacacacgttaagagatttgcgtcatgagcggaatacatatt 2054
QY 1994 tgaatgtattaga 2007
Db 2055 gatcttcacactaga 2068

RESULT 5
AAV32374/C
ID AAV32374 standard; DNA; 14455 BP.
XX AC AAV32374;
XX DT 29-SEP-1998 (first entry)
XX DE Complete sequence of the pE1/Fiber plasmid.
XX KW Circular; adenovirus type 5; pE1/Fiber plasmid; structural protein;
KW complementation; fiber protein; gene therapy; HIV; tumour; AD5;
KW early gene; Huntington's disease; Tay-Sachs disease;
KW sickle cell disease; E1 regulatory protein; ds.
XX OS Synthetic.
XX XX
XX Key location/Qualifiers
XX misc_feature 1460..4998
XX FT
XX FT /*tag= a
XX FT /*note= "AD5 E1 regulatory gene"
XX FT complement (10922..14223)
XX FT /*tag= b
XX FT /*note= "AD5 fiber gene consisting of a CMV
promoter at 5' end of this gene"
XX PD 02-APR-1998.
XX PF 24-SEP-1997; 97WO-EP05251.
XX PR 25-SEP-1996; 96US-0719806.

```

```
XX (NOV ) NOVARTIS AG.
PA (SCRI ) SCRIPPS RES INST.
PI
PI Memerow GR, Von Seggern DJ;
XX WPI; 1998-230709/20.
XX
PT Adenoviral vectors - which lack DNA encoding for structural protein
PT or fibre protein used particularly for gene therapy
XX
XX Example 1; Pages 112-131; 170pp; English.
XX
CC The present sequence is that of a pEI/Fiber plasmid used in the
CC method of the invention. The plasmid contains an adenovirus type 5 (Ad5)
CC fiber gene controlled by a CMV promoter, an Ad5 E1 gene and a pMAM
CC backbone. The invention provides adenoviral vectors having deletions of
CC all or part of various gene sequences encoding adenoviral structural
CC proteins and/or early region proteins. Deletions in these proteins would
CC allow a reduced risk of wild-type virus contamination and would also
CC allow packaging of foreign DNA in such vectors for a variety of
CC diagnostic and therapeutic applications. The adenoviral vectors having
CC deletions in the structural and/or early gene regions are produced by
CC cellular complementation of these adenoviral genes. Therefore, the
CC pEI/Fiber plasmid was used as a complementation plasmid which was
CC introduced into a host cell line where parts of the fiber and E1 gene
CC region would be stably inserted into the host cell chromosomes. The
CC resulting E1/fiber gene deficient plasmid can be used as a gene delivery
CC vector. The vectors can be used for diagnosis or gene therapy, e.g. for
CC treating conditions characterised by hyper-proliferative cells (e.g. for
CC tumors), genetic diseases (e.g. Huntington's disease, Tay-Sachs disease,
CC or sickle cell disease), or infections (e.g. HIV infection). They can
CC also be used for in vitro production of biologically active proteins.
XX
SQ Sequence 14455 BP; 3698 A; 3271 C; 3565 G; 3921 T; 0 other;
XX
Query Match 14.1%; Score 1148.2; DB 19; Length 14455;
Best Local Similarity 95.5%; Pred. No. 3e-163;
Matches 1196; Conservative 0; Mismatches 48; Indels 9; Gaps 1;
XX
QY 755 ggcgcgcctgcagatgcatactagagggccctattctatagttcacctaatactgag 814
DB 11236 GCGGCGGCTCGACATGATCTAGAGGCCCTTCTATATGTCACCTTAAATGCTAGAG 11177
QY 815 ctgcctgatacgcctgcagctgtgcttctagttgcagcagcatctgttctgcccctcc 874
DB 11176 CTGCGTGTATCAGCCTGACGTGCTCTGTATGTGCCAGCATCGTGTGGCCCTCC 11117
QY 875 ccgtgcttctctgagccctggaagtgccaccccaactgctcttctctaataaataag 934
DB 11116 CCGTGCCTTCTCTTACCCCTGGAAGGTGCCACTCCACTGCTTCTTAATAAATGAGG 11057
QY 935 aaattcatcgaatgtctgafaggtgtcatctattctcgggggggtgggggtggagc 994
DB 11056 AAATTGATCGCATTTGTCTGATGAGTGTGATCTATCTTGGGGGGGGGGGGGAGG 10997
QY 995 acagcaagggggagattgggaagacaatagcagcagcagctggtggatgctgagccta 1054
DB 10996 ACAGCAAGGGGAGATTGGGAAGACATAGCAGCATGCTGGGGAGCGGCTGGCTCTA 10937
QY 1055 tggcttctgagcggaagaagacagctgcatataagatcggccaacgcggggagag 1114
DB 10936 TGGCTTCTGAGGCGGTA-----TGCGTGTGAATATCCGACAGATGCGTAAAGAG 10886
QY 1115 cggctgcatatgggcctctcctcctcctcaactgactgactgagcctcgggtcgt 1174
DB 10885 AAATATCCCGCATAGGCGCTTCTCCGCTTCTCTGCTACTGACTCGGTGGCTGGCTGT 10826
QY 1175 tggcctgctgagcgagatcatcagctcaactcaaaagcggtataacgltaccaaagatc 1234
DB 10825 TCGGCTGCGGCGAGCGGTATACACTCACTCAAAAGGCGGTAAATGCGTTATCCACAGATC 10766
```

```
QY 1235 aggggataacgcaggaagaacatgttagcgaagaagggccagcaaaaggccaggaacgtaa 1294
DB 10765 AGGGGATACGCGACGAAAGACATGTAGACAAAGGCCAGCAAAAGGCCAGAACCTTAA 10706
QY 1295 aaaggccgctgtgctgagcttttccataggtccgcgccccctgacagacatcacaanaa 1354
DB 10705 AAAGGCCGCGTGTGCTGGCGTTTTCATAGGCTCGGCCCCCTGACAGCATCACAAAAA 10646
QY 1355 tcgacgtcaagtgcagaggtgtggcgaacccgcagacgactataaagataccagcgcttc 1414
DB 10645 TCGACGCTCAAGTCAAGAGTGGCGCAAAACCCGACAGACTATAAAGATACAGCGGTTTCC 10586
QY 1415 cccctggaagctccctgtgtgagctcctgtttccgaccctgcggttacggatacctgtc 1474
DB 10585 CCCTGGAGGCTCCCTGCTGGCGCTCTCTGTCCAGCCCTGCGCTTACCGGATACCTGTCTC 10526
QY 1475 cgccttctcccttcgggaagcgtgagccttcccaatgtcacaagcgtgatactcag 1534
DB 10525 CGCCTTCTCCCTTCGGGAAGCGTGGCGCTTTCATAGCTCAGCTGTAGTATCTCAG 10466
QY 1535 ttcggtgtaggtcgttgcctccaagctgtgctgtgtgcaagaccgccgttcagccga 1594
DB 10465 TTCGCTGTAGGTCGTTGCTCCCAAGCTGGGCTGTGTCCAGCAACCCCGTTACGCCGA 10406
QY 1595 ccgctgacctatccggtactatcgtctttagtccaaccggttaagaacgactatc 1654
DB 10405 CCGGTGCGCCTTATCCGCTAATCTATGCTGTGAATCCAAACCCGTAAGACACGACTTATC 10346
QY 1655 gccactgcaagcagccactgtgaacagatagtagagcgagatgtagcggtgtctac 1714
DB 10345 GCCACTGGCAGCAGCCACTGTGTAACAGGATTTAGCAGAGCGAGATATGAGCGGTGTCTAC 10286
QY 1715 agagttctgaagtgtgtgacctaaactacggtactaagaagacaglatgtgtatctg 1774
DB 10285 AGAGTTCTTAAAGTGTGGCGCTTAACCTACGCTACACTAAGAACGATATTGGTATCTGT 10226
QY 1775 cgcctgtgaagcaagttaccttcgcgaaagaaggtgtgtagcttgcgttcggcaaca 1834
DB 10225 CGCTCTCTGAAGCAAGTATTCCTTGGAAAAAGGTTGTAGTCTTGTGATCCGGCAACA 10166
QY 1835 aaccacgcctgtgtagcggtgtttttgtttgttgcaagcagacagatlaacgcgcaaaaaa 1894
DB 10165 AACACCGCGTGTAGCGGTGTGTTTGTGTTTGTGCAACAGCAGATTACGCCACAAAAAAA 10106
QY 1895 aggatctcaagaagaatccttgatccttcttaacggyggtctgacgctcagtgagaa 1954
DB 10105 AGGATCTCAAGAAGATCTTGTGATCTTTCATACGGGCTGTGACGCTCAGTGAAGCAAAA 10046
QY 1955 ctcaagtgtaaggatattgtlcaatgagcgatataatattgaatgtattaga 2007
DB 10045 CTCACGTTAAAGGATTTTGTTCATGAGATTATCAAAAAAGATCTTTCACCTAGA 9993
XX
XX RESULT 6
XX AAA59050/C
XX ID AAA59050 standard; DNA; 14455 BP.
XX
XX AAA59050;
XX AC
XX XX
XX 07-NOV-2000 (first entry)
XX
XX Nucleotide sequence of the E1/fiber-expressing plasmid pEI/Fiber.
XX Adenovirus: tripartite leader; adenovirus vector particle; gene delivery;
XX E1 gene; fiber gene; ss.
XX
XX Synthetic.
XX Adenovirus type 5.
XX
XX WO200042208-A1.
XX
XX PD 20-JUL-2000.
XX
```

PF 14-JAN-2000; 2000WO-EP00265.
XX
PR 14-JAN-1999; 99US-0115920.
XX
PA (NOVS) NOVARTIS AG.
PA (NOVS) NOVARTIS-ERFINDUNGEN VERN GES MBH.
PA (SCRI) SCRIPPS RES INST.
XX
PI Nemerow GR, Von Seggern DJ, Hallenbeck PL, Stevenson SC;
PI Skripchenko Y;
XX
DR WPI: 2000-4/6068/41.
XX
PT New nucleic acid comprising an adenovirus tripartite leader nucleotide
PT for producing high-capacity and targeted vectors for adenovirus-based
PT gene therapy -
XX
PS Example 1; Page 160-164; 212pp; English.
XX
CC The specification describes a nucleic acid molecule comprising an
CC adenovirus (AV) tripartite leader (TPL) nucleotide with a sequence
CC comprising two different TPL exons or three same or different TPL
CC exons. The nucleic acid is used to produce an adenovirus vector particle,
CC deliver an exogenous gene to a target cell, pseudotype recombinant viral
CC vectors, target an adenovirus vector to a cell, produce a modified
CC adenovirus, deliver a heterologous gene to an animal and produce a
CC glues adenoviral vector particle. The present sequence represents
CC a El/fiber-expressing plasmid, which was used for complementation of
CC El/fiber-gene deleted adenoviruses.
XX
SQ Sequence 14455 BP; 3698 A; 3271 C; 3565 G; 3921 T; 0 other;

Query Match 14.1%; Score 1148.2; DB 21; Length 14455;
Best Local Similarity 95.5%; Pred. NO. 3e-163;
Matches 1196; Conservative 0; Mismatches 48; Indels 9; Gaps 1;
QY 755 gggcgccgctcagcatcgtatcgtagagggccctatctatagtgccacctaagctagag 814
DB 11236 GCGGCGCGTCAGCATGATCATAGAGGCCCATCTTATAGTCACCTAAATGCTAGAG 11177
QY 815 ctgcgtgcatcagcctcgacgtgctctcctagtgtccagccatcgtgtgttgcctcc 874
DB 11176 CTCGCTGATCAGCCCTCGACTGCTGCTTGTGTCAGCCACATCTGTGTTGCCCTCCC 11117
QY 875 ccgtgacctcttgaccttggaaggtgacccactgcctcttccataaataatgag 934
DB 11116 CCGTGCCCTTCTGACCTTGGAAGGTGCCACTCCACATGCTCTTCTTAATAAATGAAG 11057
QY 935 aaatgcacatcgtctgtgtagtgatgtcatctatctctgagggtggtggtggtgag 994
DB 11056 AAATTGCAATGCGATGTCGTGAGTAGTGTCTATCTATCTGSGGGGTGGGGGCAAG 10997
QY 995 acagcaaggggagagatgtgggaagacaatagcaagcagtcgtggtggtggtccta 1054
DB 10996 ACAGCAAGGGGAGGATGGAGACAAATAGCAGCATGTGGGATCGGGGCTCTA 10937
QY 1055 tggcttcgaaagcgaaagacacagctgcatatgaatcgcccaagcggtgggaagag 1114
DB 10936 TGGCTTCGAGAGCGGTA-----TGGCGGTGTAATAATCCGCACAAATGCGTAAGAG 10886
QY 1115 cggttggtatgtgggagcctcttcgcctctcgtctcactgactgctgagcgtgagtcgt 1174
DB 10885 AAATATACCGCATCAGGCGCTCTTCGCTTCTCCTGCTCATGACTGCTGCGTCCGTCGT 10826
QY 1175 tggcgctgcgagcgatcagcatcactcaaaagcggttaatacgttatccacaagaatc 1234
DB 10825 TCGGCTGCGGAGGAGGTATCATGCTCACTCAAAAGCGGTAAATACGTTATCCAGAAATC 10766
QY 1235 aggggataacgcaggaagaacatgtgagcaaaagcgagcaaaagcgcaaggaacgta 1294
DB 10765 AGGGATTAACGACAGGAAGAACATGTGAGCAAAAGGCCACGCAAAAGCGAGAACCGTA 10706

QY 1295 aaagcgcgctgtcgtgagcttttccatagctccgcccccctgacagatcacacaanaa 1354
DB 10705 AAAGCCCGGTCGTCGGCGTTTTCATATAGGCTCCGCCCCCTGAGAGCATCACAAAAA 10646
QY 1355 tgcagctcaagtcagaagtgagcaaaaccgacaggaactataaagatacaggcttcc 1414
DB 10645 TCGAGCTCAAGTCAGAGATGGCGAAACCCGACAGACTATTAATACACAGCGCTTTC 10586
QY 1415 cccttgaaagctccctcgtgagctctcgtcttcagacccctgcgcgttaaccgatactg 1474
DB 10585 CCTGGAAGCTCCTCTGTCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 10526
QY 1475 cgctctcctccctcgggaagcggtgagcttctcctaagtgacagcgtgtagtlatccag 1534
DB 10525 GCGCTTTCCTCCCTTGCGGAGCGGTGGCGCTTCTCATAGCTCACGCTGATAGTCTCAG 10466
QY 1535 ttcggtgtagtgcgttcgctcgaagtggtgctgtgtgcagaaccccccggttcagccga 1594
DB 10465 TTCGGTGTAGGTGCTTCGCTCCAAAGCTGGGCTGTGTGCAAGAACCCCGCTTCAGCCGA 10406
QY 1595 ccgctgcgcttattccggttaactatcgtcttgagtcgaaccccggttaagacagacttacc 1654
DB 10405 CCGTGCCTTATCCGGTAACTATGCTTGAAGTCCAAACCGGTAAAGACAGACTTATC 10346
QY 1655 ggcactggcagcagccactggttaacagatlagcagagcgagtagttagcggtgtac 1714
DB 10345 GCCATGGGAGCAGCAGCATCGTAAACAGAGTTAGCAGAGGAGAGTATGTCGGGCTAC 10286
QY 1715 agagttctgaaagtgtgagcctaactaagcgctaacactagaagaagatattgtgatactg 1774
DB 10285 AGAGTTCCTTGAAGTGTGTGCCCTTAACCTACGCTACACTAAGAGACAGATTTTGTATCTG 10226
QY 1775 cgctgtgctgaagccagttactctcgtgaaagaagatgtgtagctctgtatccgcgaaca 1834
DB 10225 CCTCTGCTGAAAGCCAGTTACTCTTCGAAAAAGAGTGTGATCTTGAATCCGCGCAACA 10166
QY 1835 aaccacgcgtgtgtagcgtgtgttttltgttgcagcagcagatlaacgcgcagaaaaa 1894
DB 10165 AACCAACCGGTGTAGCGGAGGTTTTTTTGTGCAAGCAGCAGATTAACCGCAGAAAAA 10106
QY 1895 aggatctcaagaagatccttgaatcttctcaggggtgtgagcgtcagtggaagaaaa 1954
DB 10105 AGCATCTCAAGAAAGATCTTTTATCTTTTTCACGGGGTCTGACGCTCACTGGAAGAAAA 10046
QY 1955 ctcaagttgaaggttttgatgcatgagcgcgataatttgatattaga 2007
DB 10045 CTCACGTTAAGGAGTTTGTGTCATGAGATTATCAAAAGATCTTCACTTACA 9993
RESULT 7
ID AAA53872 standard; DNA; 5377 BP.
XX
AC AAA53872;
XX
DT 03-JAN-2001 (first entry)
XX
XX Expression vector pRIG14.
XX
KW Vector; endogenous gene; activation; over-expression;
KW erythropoietin; growth hormone; drug discovery;
KW granulocyte colony stimulating factor; ds.
OS Synthetic.
XX
PN WO200049162-A2.
XX
PD 24-AUG-2000.
XX
PF 22-FEB-2000; 2000WO-US04429.
XX
PR 19-FEB-1999; 99US-0253022.
PR 08-MAR-1999; 99US-0263814.

PR 26-MAR-1999: 99US-0276820.
XX (ATHE-) ATHERSYS INC.
XX Harrington JJ, Sherf B, Rundlett S;
PI MPI: 2000-549276/50.
XX
XX Non-targeted activation of endogenous genes, e.g. for the production of
PT erythropoietin, growth hormone or granulocyte-colony stimulating factor
PT proteins and for drug discovery
PS Disclosure: Fig 29: 240pp: English.
XX
XX New methods, vectors and cells are described for non-targeted
CC activation and over-expression of endogenous genes. The vector
CC constructs comprise transcriptional regulatory sequences (TRS) and
CC unpaired splice donor sequences (USDS), preferably the vectors
CC comprise (in sequential order) a TRS, an USDS, a rare cutting
CC restriction site (RCRS) and a linearization site (LS) with a second
CC TRS linked to a selectable marker (SM) lacking a polyadenylation
CC signal. The methods, vectors and cells comprising the vectors may
CC be used for the non-targeted activation and over-expression of
CC endogenous genes, e.g. for the production of proteins (including
CC erythropoietin, growth hormone or granulocyte-colony stimulating
CC factor) and drug discovery. The advantage of these methods are that
CC endogenous genes including those associated with human disease and
CC development, may be activated and isolated without prior knowledge
CC of the sequence structure, function or expression profile of
CC the genes being known.
XX
XX Sequence 5377 BP; 1248 A; 1465 C; 1437 G; 1225 T; 2 other:
SQ

Query Match 14.0%; Score 1140; DB 21; Length 5377;
Best Local Similarity 74.3%; Pred. No. 5.7e-162;
Matches 1684; Conservative 0; Mismatches 300; Indels 284; Gaps 8;

QY 23 gttgacattgatatgacagtgatattataatgaatcaatcaatgaagggtcattgata 82
DB 2058 gtgtgacattgatatgacagtgatattataatgaatcaatcaatgaagggtcattgata 2117
QY 83 gccacataatgagaggttcgcggttaccataacttcggttaactgcccgcggtgcagccg 142
DB 2118 gccacataatgagaggttcgcggttaccataacttcggttaactgcccgcggtgcagccg 2177
QY 143 ccaacgaccccgcccgccatgacgtcaataatgacgtatgtcccatgtaacgccaatag 202
DB 2178 ccaacgaccccgcccgccatgacgtcaataatgacgtatgtcccatgtaacgccaatag 2237
QY 203 ggaacttcacatgacgtcaatggtgacatttaacgtaaacgtgccacttgagctac 262
DB 2238 ggaacttcacatgacgtcaatggtgacatttaacgtaaacgtgccacttgagctac 2297
QY 263 ataagtgtatcatatgccaagtaagcccgctattgacgtcaatgaggttaactgcccgcg 322
DB 2298 ataagtgtatcatatgccaagtaagcccgctattgacgtcaatgaggttaactgcccgcg 2357
QY 323 cctcgatattgcccagtaataacttaacttaactggaacttccacttggcaatacactacg 382
DB 2358 cctcgatattgcccagtaataacttaacttgaacttccacttggcaatacactacg 2417
QY 383 taatagtcacgtcattacacatggtgacgtgtttggcagtaacatcattggcggtgat 442
DB 2418 taatagtcacgtcattacacatggtgacgtgtttggcagtaacatcattggcggtgat 2477
QY 443 agcggtttgactacgaggtttccaaagtctccaccccatgacgttaaatggaggttgt 502
DB 2478 agcggt-ttgcactacgaggtttccaaagtctccaccccatgacgttaaatggaggttgt 2536
QY 503 ttggagcccaaaatcaacgagactttccaaatgtctaaact-----cgccccc 554
DB 2537 ttggagcccaaaatcaacgagactttccaaatgtctaaactcgatcgccccc 2596

QY 555 attgacgcaaatggcggttagcggtatgacgtgtgaggtctatatgaagcagagctctg 614
DB 2597 gttgacgcaaatggcggttagcggtatgacgtgtgaggtctatatgaagcagagctctt 2656
QY 615 gtaactagagaacccaact-----gtaactagatc 648
DB 2657 agtgaaccgtcagatacactagaattatgctgaggttagttatacagtaaatgtcctaac 2716
QY 649 gaaatatacgtacactatagggagacccaagcttggtaacgagctcgagctcagactag 708
DB 2717 gcagtcagtgcttcgacacacagctcgaacttaagctgcagctgctctccttaataa 2776
QY 709 taacggcgcaggtgt----- 725
DB 2777 ccacgctaacaggtgagtaactcgatctgtacctaagaagagcctatctgcccagtta 2836
QY 726 ----- 725
DB 2837 gcagtcgaagaagaatttaagaagccgaacacgctcattgagcccggaagtgcga 2896
QY 726 -----ctggaatttcgagatatacacaacttgcgccc 760
DB 2897 gcccgatctcccatcggtgattgctgcgcatatagcgccagcaacgcgacacttgcg 2956
QY 761 gctcagacatgcatctgtagggccctattctatagtgtcaacttaatgctagagctgc- 819
DB 2957 ccggtgtgcgagcagatgctgcggtcagtgagatccacagaggtgtgtgtgcg 3016
QY 820 ----tgacagcctcagactgtgctctcagttgctgcaagcactgtgttgcctccccc 876
DB 3017 atgacgtcgtagtcgataagtggtcccaagttagcgaagcagagatgtggtgcggtcga 3076
QY 877 gtgcctccttga----- 889
DB 3077 aagcggtcgcagcaggtgtctccgagacggtgtgcatagaattgcataacgcatatagc 3136
QY 890 -----ccctggaaaggtgccactccactgtccttctcta 923
DB 3137 gctagatccttgcagtagctgtaggcgcgacccgcggtggaagcttcacattgttccct 3196
QY 924 ataaatgagaaatgtcatcgcatgtctcgtagtaggtgtcatctatctctggtggtg 978
DB 3197 tagtgaaggttaatttcgagcttgcggttaactatggtcatalagctgttctctgtgtgaa 3256
QY 979 -----gttgggtgtgggtcaggaagcaggaaggggag 1009
DB 3257 ttgtatcgcgtcacaaatccacaacatacgaagcgggaagcataaagtgttaagcctggg 3316
QY 1010 attgggaagaacaatagcagatgctgtggagtgcgtgtggtctatgtgtctt----- 1062
DB 3317 gtgcctatagtgtagcttaactcacatlaattgcgtgtgcctcacttcccgcttccagt 3376
QY 1063 ----gagggcgaagaagaacacgtgcatatgaatcggcgaacggtgggagaggggtt 1119
DB 3377 cgggaaacctgtcgttcgacagctgcatatgaatcggcgaacggtgggagaggggtt 3436
QY 1120 ttgctattggcgctcttccgctctcctcgtcctcactgactgactgctgtgcgtctgcgcg 1179
DB 3437 ttgctattggcgctcttccgctctcctcgtcctcactgactgactgctgtgcgtctgcgcg 3496
QY 1180 ttgcggcagcggtatcagctcactcaaaagcggtataaggttatccagaatcagggg 1239
DB 3497 ttgcggcagcggtatcagctcactcaaaagcggtataaggttatccagaatcagggg 3556
QY 1240 ataacgagaagaagaatgtgagcaaaagcggaagggccaggaacggttaaaagg 1299
DB 3557 ataacgagaagaagaatgtgagcaaaagcggaagggccaggaacggttaaaagg 3616
QY 1300 ccggttgcgtgcgtttttccataggtctccgcccccttgcagagcatcaacaaatcgac 1359
DB 3617 ccggttgcgtgcgtttttccataggtctccgcccccttgcagagcatcaacaaatcgac 3676

QY 615 gctaaatagagagaccactgtctactgtgcttactcgaaatlaaactactactatagga 674
 Db 739 agtgaacccgtacagatcacactgaaattcttgac---gacctactgtactaaacggcatagagcc 795
 QY 675 gaccaacagcttgttaccagactcgcagatccactagtaacgagccgacggtgtgttggaattcc 734
 Db 766 tccctgcagatcaactaagaagcttatttgcgttagttatcaacgtaaatgttgaacgag 855
 QY 735 tgcagatatccatacacttgcgcgcgtcgcagcatgcatctagagggccattctata 794
 Db 856 tcaagctcttgcacacacacagatcgaacttaagctgcagtgactctcttaataccacca 915
 QY 795 gtgtacactaatatgttaagctcgtgtgtatcgacctgcagctgtgcttctagtgtccagcc 854
 Db 916 tggctacaggtgtgagtaactcgcacttaagagagggcctatctgcgcagtttagcagtcgaa 975
 QY 855 atctgttgttgcacctccctcccgctccttcccttgacccttgaaaggtgcacactccactgt 914
 Db 976 gaagaaagttttaagagagccgaaacaaagcgtctcagtagcccgaaagtgtgcagccgaltct 1035
 QY 915 ccttccctaataaaatgaggaatgtcatgcgcatgtctgtgagtagtgtcattctatct 974
 Db 1036 tcccatcggtgtatgtcgcgcatatagcgcagcaacccgacctgtgtgcgc----- 1088
 QY 975 ggggggtgtgggtgtggcgacagcaagggaggtatgttgagaagcaatagcagcatgc 1034
 Db 1089 -----ggctgtatgcgcgcgcacgaatgtctcggcgttagagagatccacagacgtgtgt 1139
 QY 1035 tggggtatgcggtgtggcctctatgtgcttcgagggcgaaagcaacgtcatatgaatc 1094
 Db 1140 ggtcgcacatgtat-----cgctgtatgtcatgtgtg 1168
 QY 1095 ggcaacgcgcggggagagagcggtttgcgtatgtggcgctcttccgcttccctcact 1154
 Db 1169 ctcaacatagcagaagcgagcagactgtggcgcgcaaaagcggtcgcgacagtgctccga 1228
 QY 1155 gactcgcgtcgcgtcgtgtcgttcgctgcgtgcgagcggtatcagctccactcaaaagcggt 1214
 Db 1229 gaacgggtgtgcgcatagaatgtcatcaacgatatagcgtatagcttctgtcagatgtga 1288
 QY 1215 atacggtatccacagaatcaggggatacgcaggaagaacatgttgagcaaaagcgag 1274
 Db 1289 gatctgtcagac-----catgtgagcaaaagcgag 1319
 QY 1275 caaaagccagagaaacgtaaaaagcgcggttgcgtgttcttccatagctccgccc 1334
 Db 1320 caaaagccagagaaacgtaaaaagcgcggttgcgtgttcttccatagctccgccc 1379
 QY 1335 cctgaagagcatcacaaaatcgcagctcaagtcaaggtgtggcgaacccgagagacta 1394
 Db 1380 cctgaagagcatcacaaaatcgcagctcaagtcaaggtgtggcgaacccgagagacta 1439
 QY 1395 taagagataccagagcggttcccccctggaaagctccctcgtgcgttccctgtccgacctg 1454
 Db 1440 taagagataccagagcggttcccccctggaaagctcccccgtgcgttccctgtccgacctg 1499
 QY 1455 ccggttaacgagatacctgtccgcttcttccctccttcggaagcggtgcgcttctcaatgc 1514
 Db 1500 ccggttaacgagatacctgtccgcttcttccctccttcggaagcggtgcgcttctcaatgc 1559
 QY 1515 tcaagctgtatgtatcagttcagttgtagtgcgttcgctccaaagctgtggtgtgac 1574
 Db 1560 tcaagctgtatgtatcagttcagttgtagtgcgttcgctccaaagctgtggtgtgac 1619
 QY 1575 gaacccccgttgcagccgagacgcgcgttccctatccgtaactatcgctctgagctcaac 1634
 Db 1620 gaacccccgttgcagccgagacgcgcgttccctatccgtaactatcgctctgagctcaac 1679
 QY 1635 ccggttaacgagactatcgcgaactgtgcagcagcagcactgtgtaaacagattagcagagcg 1694
 Db 1680 ccggttaacgagactatcgcgaactgtgcgaactgtgcagcagcactgtgtaaacagattagcagagcg 1739

QY 1695 aggtatgtagcgcgtgtcactaagagttcttgaagtgtgtgcccactaagcgtacactaga 1754
 Db 1740 aggtatgtagcgcgtgtcactaagagttcttgaagtgtgtgcccactaagcgtacactaga 1799
 QY 1755 aggaacagatattgtgtatctgtcgtctgtcgttgaagcagttaccttcggaaaaagattgt 1814
 Db 1800 aggaacagatattgtgtatctgtcgtctgtcgttgaagcagttaccttcggaaaaagattgt 1859
 QY 1815 agctcttgatccgcgcaaaacacccgcgtgtgtagcgtgtgtttttttgtttgtgaagcag 1874
 Db 1860 agctcttgatccgcgcaaaacacccgcgtgtgtagcgtgtgtttttttgtttgtgaagcag 1919
 QY 1875 cagattacgcgcagagaaaaaagatcctcagaagaagatccttactatcttctacgggtct 1934
 Db 1920 cagattacgcgcagagaaaaaagatcctcagaagaagatccttactatcttctacgggtct 1979
 QY 1935 gacgtcaggtggaacgaacaaactcacggttaagagatttgcatagcgcgatacatatt 1994
 Db 1980 gacgtcaggtggaacgaacaaactcacggttaagagatttgcatagcatacaaaaag 2039
 QY 1995 gaattgattaga 2007
 Db 2040 atcttcacactaga 2052

 RESULT 9
 AAV38297
 ID AAV38297 standard; DNA; 5446 BP.
 XX
 AC AAV38297;
 DT 26-OCT-1998 (first entry)
 XX
 DE Plasmid pCDNA3.
 XX
 KW Plasmid pCDNA3; pneumococcal surface protein A; PspA; infection;
 KW Streptococcus pneumoniae; sepsis; otitis media; meningitis;
 KW bacteremia; pneumonia; vaccine; genetic immunisation; ss.
 XX
 OS Chimeric - human cytomegalovirus.
 XX
 PN MO9824927-A1.
 XX
 PD 11-JUN-1998.
 XX
 PF 04-DEC-1997; 97WO-US22847.
 XX
 PR 04-DEC-1996; 96US-0759505.
 XX
 PA (UYAL-) UNIV ALABAMA.
 XX
 PI Briles DE, Curriel DT, McDaniel LS;
 XX
 DR WPI: 1998-333343/29.
 XX
 PT Plasmid containing pneumococcal epitope for expression in eukaryotic
 PR cells - useful for eliciting immunological response to pneumococcal
 PT infection or sepsis
 XX
 PS Example 1; Fig 1B1 to 1B-5; 47pp; English.
 XX
 CC This is the DNA sequence of plasmid pCDNA3 (Invitrogen). A portion
 CC of the gene (see AAV38298) that codes for respiratory syncytial virus
 CC glycoprotein G (RSV) has been amplified, digested with KpnI and
 CC ligated into KpnI-digested pCDNA3 upstream of the multiple cloning
 CC site of pCDNA3 and downstream of the human cytomegalovirus
 CC immediate early (HCMV-IE) promoter to create pCF41. A full-length
 CC coding sequence of Streptococcus pneumoniae Rxi pneumococcal
 CC surface protein A (PspA) was then inserted into pCF41 to create
 CC a fusion between rsvG and PspA. Intramuscular immunisation of
 CC BAB/c mice with the resulting plasmid, designated PRSD2601,
 CC induced protection against an otherwise lethal challenge with a
 CC capsular type 3 pneumococcus. A plasmid for expression of

CC pneumococcal epitope DNA in eukaryotic cells is claimed. The
CC plasmid includes a promoter for driving expression in a eukaryotic
CC cell (e.g. HCMV-IE), DNA encoding a leader sequence (e.g. of RSVG)
CC which facilitates expression, translation through or transport of
CC the expression product in a eukaryotic cell membrane, and DNA
CC encoding a pneumococcal epitope such as PSPA. The invention also
CC provides a vaccine comprising the plasmid and a suitable carrier or
CC diluent, and optionally one or more cytokines or DNA encoding them,
CC or a bacterial delivery system. The vaccine is used to elicit an
CC immunological response in a host, including humans, susceptible to
CC pneumococcal infection or sepsis. The plasmid can also be used to
CC express a pneumococcal epitope of interest in vitro.

XX
XX
Sequence 5446 BP; 1255 A; 1417 C; 1390 G; 1384 T; 0 other;

Query Match 13.3%; Score 1082; DB 19; Length 5446;
Best Local Similarity 100.0%; Pred. No. 2.5e-153;
Matches 1082; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gatgacggcgccagatagacgcttgacattgattatgacagtattatgaatca 60
DB 210 gatgacggcgccagatagacgcttgacattgattatgacagtattatgaatca 269
QY 61 attacggggatcattagttcattatgacccatataatgagttccggttacataatcagta 120
DB 270 attacggggatcattagttcattatgacccatataatgagttccggttacataatcagta 329
QY 121 aatggccgcctgctgtgacgcccacgaaccccgccattgacgttcaataatgacgtat 180
DB 330 aatggccgcctgctgtgacgcccacgaaccccgccattgacgttcaataatgacgtat 389
QY 181 gtcccatagtaacgccaatagggacttccattgacgccaatggtgtgacattatcag 240
DB 330 gtcccatagtaacgccaatagggacttccattgacgccaatggtgtgacattatcag 449
QY 241 taaactgcccacttggcagtatcaatagttatcatabatgccaagtacgccccattgac 300
DB 450 taaactgcccacttggcagtatcaatagttatcatabatgccaagtacgccccattgac 509
QY 301 gtcaatgacggttaaatgcccgccttggcattatgcccataatgacctatgagacctt 360
DB 510 gtcaatgacggttaaatgcccgccttggcattatgcccataatgacctatgagacctt 569
QY 361 cctacttggcagtatcattacgattatgacgcgctattacacatggtgacggtttcg 420
DB 570 cctacttggcagtatcattacgattatgacgcgctattacacatggtgacggtttcg 629
QY 421 cagtaatacaatggcggtgtagcggtttgactcaacggggtattccaaagtcacccc 480
DB 630 cagtaatacaatggcggtgtagcggtttgactcaacggggtattccaaagtcacccc 689
QY 481 attgacgtcaatggggaggtttgttttggcaccaaatcaacggggtattccaaatgtcgt 540
DB 690 attgacgtcaatggggaggtttgttttggcaccaaatcaacggggtattccaaatgtcgt 749
QY 541 aacaactcgcccccattggacgaatggtgagcggtgtagcggttgaagtgggaggtctatata 600
DB 750 aacaactcgcccccattggacgaatggtgagcggtgtagcggttgaagtgggaggtctatata 809
QY 601 agcagagctctctgcttaactagagaaccactgtctactgtgcttgcgaatataacg 660
DB 810 agcagagctctctgcttaactagagaaccactgtctactgtgcttgcgaatataacg 869
QY 661 actactataggagagaccagcttggtagcgagctcgatccactagtataagggcgcca 720
DB 870 actactataggagagaccagcttggtagcgagctcgatccactagtataagggcgcca 929
QY 721 ggtgtgtgaattcttcagatatccatcacactggcgcgccgctcgagacgactctagag 780
DB 930 ggtgtgtgaattcttcagatatccatcacactggcgcgccgctcgagacgactctagag 989
QY 781 ggcctcatctatagtgatccctaataatgtagcgctcgatcagcgctcgtactgtgct 840

DB 990 ggcctcatctatagtgatccctaataatgtagcgctcgatcagcgctcgtactgtgct 1049
QY 841 tctattgcccagccatctgtgttggccctcccccgcgttcccttgaacctggaaggt 900
DB 1050 tctattgcccagccatctgtgttggccctcccccgcgttcccttgaacctggaaggt 1109
QY 901 gccactcccatctgtcttctcctaataaataatgaggaattgcatctgtgtgtagag 960
DB 1110 gccactcccatctgtcttctcctaataaataatgaggaattgcatctgtgtgtagag 1169
QY 961 tctcatctattcttgagggttgagggtgaggaacgaaggggagagattggaagac 1020
DB 1170 tctcatctattcttgagggttgagggtgaggaacgaaggggagagattggaagac 1229
QY 1021 aatagcagggcatgcttgaggatcggttggtctatagttcgttgcagggcgaaccagc 1080
DB 1230 aatagcagggcatgcttgaggatcggttggtctatagttcgttgcagggcgaaccagc 1289
QY 1081 tg 1082
DB 1290 tg 1291

RESULT 10

AAA53871
ID AAA53871 standard; DNA: 5245 BP.

AAA53871;

03-JAN-2001 (first entry)

Expression vector pRIG22b.

Vector: endogenous gene; activation; over-expression;

erythropoietin; growth hormone; drug discovery;

granulocyte colony stimulating factor; ds.

Synthetic.

WO200049162-A2.

22-FEB-2000; 2000WO-US04429.

19-FEB-1999; 99US-0253022.

08-MAR-1999; 99US-0263814.

26-MAR-1999; 99US-0276820.

(ATHE-) ATERSYS INC.

Harrington JJ, Sherf B, Rundlett S;

WPI; 2000-549276/50.

Non-targeted activation of endogenous genes, e.g. for the production of

erythropoietin, growth hormone or granulocyte-colony stimulating factor

proteins and for drug discovery

Disclosure; Fig 16; 240pp. English.

New methods, vectors and cells are described for non-targeted

activation and over-expression of endogenous genes. The vector

constructs comprise transcriptional regulatory sequences (TRS) and

unpaired splice donor sequences (USDS), preferably the vectors

comprise (in sequential order) a TRS, an USDS, a rare cutting

restriction site (KCRS) and a linearization site (LS) with a second

TRS linked to a selectable marker (SM) lacking a polyadenylation

signal. The methods, vectors and cells comprising the vectors may

be used for the non-targeted activation and over-expression of

endogenous genes, e.g. for the production of proteins (including

erythropoietin, growth hormone or granulocyte-colony stimulating

Db 1987 gatcttcaccctaga 2000

RESULT 11
AAV40007
ID AAV40007 standard; DNA: 4026 BP.

AAV40007:
15-FEB-1999 (first entry)

Plasmid pCTM1.

E2F; transcription factor; human; retinoblastoma protein RB;
bladder cancer; restenosis; angioplasty; diabetic retinopathy;
thyroid hyperplasia; Grave's disease; psoriasis;
benign prostatic hypertrophy; Li-Fraumeni syndrome;
peripheral vascular disease; therapy; plasmid pCTM1; ss.

OS Chimeric - cytomegalovirus.
OS Chimeric - mastadenovirus.
OS Chimeric - bacteriophage T7.
OS Chimeric - bacteriophage SP6.
OS Chimeric - rhesus macaque polyoma virus.
OS Chimeric - Bos taurus.

Key Location/Qualifiers
FT promoter 209..864 /tag= a
FT misc-feature 907..1074 /note= "CMV promoter"

FT intron 1075..1253 /tag= b
FT intron 1305..1322 /function= tripartite leader sequence
FT promoter 1305..1322 /note= "hybrid SV40 late intron"

FT misc-feature 1851..4026 /tag= d
FT misc-feature 1851..4026 /note= "SP6 promoter"

FT CDS /tag= e
FT CDS complement (3032..3890) /note= "pUC19 backbone H3 to AatII"

FT CDS /tag= f
FT CDS complement (3032..3890) /note= "AMP-ORF"

PN WO9821228-A1.
PD 22-MAY-1998.
PF 13-NOV-1997; 97MO-US21821.
PR 14-FEB-1997; 97US-0801092.
PR 15-NOV-1996; 96US-0751517.
PA (CANJ-) CANJ1 INC.
PI Antelman D, Gregory RJ, Wills KN;
DR WPI; 1998-297858/26.
XX New fusion polypeptide of, e.g. transcription factor - used to
XX treat, e.g. hyper-proliferative disease such as cancer and
XX restenosis
XX Example 1; Fig 6; 91pp; English.
XX This is the nucleotide sequence of pCTM1, a plasmid that was
XX constructed from pCTM (see AAV40006) by digesting pCTM with XhoI and
XX NotI and subcloning a 180 bp intron XhoI-NotI fragment from a
XX pCMV-beta-gal vector. Plasmid pCTM1 has been used as a vector for
XX the expression of fusion proteins of the invention that comprise

CC retinoblastoma protein (BP, see AAV62465) and E2F transcription
CC factor (see AAV62464). Such fusion proteins, particularly expressed
CC from gene therapy vectors, are used to treat hyperproliferative
CC conditions, specifically cancer (particularly of the bladder) or
CC restenosis. They are more effective in repressing transcription of
CC the E2F promoter than RB alone and cause cell-cycle arrest in a
CC variety of cells.
XX
SQ Sequence 4026 BP; 978 A; 1021 C; 982 G; 1045 T; 0 other;

Query Match 13.2%; Score 1070.2; DB 19; Length 4026;
Best Local Similarity 69.7%; Pred. No. 1.5e-151;
Matches 1919; Conservative 0; Mismatches 88; Indels 746; Gaps 5;

QY 1 gatgacggcgacagatatacgcgttgatcatgtatattgactagttatataagtaacaa 60
Db 210 gatgacggcgacagatatacgcgttgatcatgtatattgactagttatataagtaacaa 269

QY 61 attacggggtcattagttcattatagcccatatatagttcggttaacaaacttaagta 120
Db 270 attacggggtcattagttcattatagcccatatatagttcggttaacaaacttaagta 329

QY 121 aatggcccgctgctgtagccgccaacgaccccgccattgacgttaataatagcgtat 180
Db 330 aatggcccgctgctgtagccgccaacgaccccgccattgacgttaataatagcgtat 389

QY 181 gtcccatagtaacgcacataggagactttccattgacgttaacgttgatattatagc 240
Db 390 gtcccatagtaacgcacataggagactttccattgacgttaacgttgatattatagc 449

QY 241 taactgccacattgacgtatcatcaatgaatgataatagccaagtacgccccattatgac 300
Db 450 taactgccacattgacgtatcatcaatgaatgataatagccaagtacgccccattatgac 509

QY 301 gtcaatgacggttaatggccgctgcatatagcccatatgacattatggagactt 360
Db 510 gtcaatgacggttaatggccgctgcatatagcccatatgacattatggagactt 569

QY 361 cctacttgacgtatcatatagcgtatgataatgataatgataatgataatgataatgata 420
Db 570 cctacttgacgtatcatatagcgtatgataatgataatgataatgataatgataatgata 629

QY 421 cagtacatcaatggcggtgtagatagcgtttgactcaacgggagattccaagttccacccc 480
Db 630 cagtacatcaatggcggtgtagatagcgtttgactcaacgggagattccaagttccacccc 689

QY 481 attgacgtcaatggaggtttgttttgacccaataacgggagattccaagttccacccc 540
Db 690 attgacgtcaatggaggtttgttttgacccaataacgggagattccaagttccacccc 749

QY 541 aacaactccgcccatattacgcaaatggcggttagcggtgacgttgaggaggtctatata 600
Db 750 aacaactccgcccatattacgcaaatggcggttagcggtgacgttgaggaggtctatata 809

QY 601 agcagagctctctggttaactagagaacccaactgcttactggtctatagaatataacg 660
Db 810 agcagagctctctggttaactagagaacccaactgcttactggtctatagaatataacg 869

QY 661 actcaactataggagagaccgaactgg----- 687
Db 870 actcaactataggagagaccgaactggcggtgtacccaactctctccgcatcgctgtctg 929

QY 688 ----- 687
Db 930 cgaggcgacgtgttggtgctgcggttgaggaacaaactcttcggtcttccagttactc 989

QY 688 ----- 687
Db 990 ttgagtcggaacccgctgcctccgaacggttactcgcacacggaggacctgagcgt 1049

QY 688 ----- 687

Db 1050 ccgcatccgacggatctcgaaaacctctcgaggaactgaaaaaccagaaatgtaactgtga 1109
QY 688 -----taccgagctcggaatccactgtaacgcccgcgtg 723
Db 1110 agtttagctcttctgtcttttatttcaagtcgccgataccggtggtggtgcaatacaag 1169
QY 724 tgcgtgaattctcgagatat----- 743
Db 1170 aactgcccctcagtgagtggtgttacttcttaggcctgtaaggaagtgtactctgc 1229
QY 744 -----ccatacaactggtggccgc-----tcgagatgcat 774
Db 1230 tctaaagctcggaattgtacccgcgctgcagcttagacgaatttcgctgacgata 1289
QY 775 ctgagagggcccttcttatatgtgtgcaactaaatgtcagctgcgtatcaagcctgact 834
Db 1290 tccatgggcccattctcatatgctgcaactaaatgtcagagctgcgtatcaagcctgact 1349
QY 835 gtgcctctagttgcaagcactgtgtgttggccctcccccgtgcttcccttgaccctg 894
Db 1350 gtgcctctagttgcaagcactgtgtgttggccctcccccgtgcttcccttgaccctg 1409
QY 895 gaaggtgcaccccccactgtccttctcctaaataaatagaggaatgtcgcattgtctg 954
Db 1410 gaaggtgcaccccccactgtccttctcctaaataaatagaggaatgtcgcattgtctg 1469
QY 955 agtaggtgtcattctctatctctgggggtgtgggtgtgggcaagcaagggggagagattgg 1014
Db 1470 agtaggtgtcattctctatctctgggggtgtgggtgtgggcaagcaagggggagagattgg 1529
QY 1015 gaagacaatagcaggaactgtctgggga----- 1040
Db 1530 gaagacaatagcaggaatgtcagcaaccaagcgcaacctgcatacagagatttcga 1589
QY 1041 ----- 1040
Db 1590 ttccaccgcgcctctctatgaagaagtgtggcttcggaatcgtttccgggaagcgcgctg 1649
QY 1041 ----- 1040
Db 1650 gatgatccctcagcgcggtatcctatgtcgtgtctcgcacccccaactgtttat 1709
QY 1041 ----- 1040
Db 1710 tgcagcttaatagtgttaacaataaagaatgacatcaacaatttcaacaataaagaatt 1769
QY 1041 ----- 1040
Db 1770 ttttcaatgcatcttagtgtgtgtgttgcataaactcataatgatatcatatgtctg 1829
QY 1041 ----- 1040
Db 1830 tatacgttcgacctctagctagagctgtgctgatacatgtctatagctgttccgtgtg 1889
QY 1041 ----- 1040
Db 1890 aaattgttatccgctcaacaattccacaacaatacgaagccggaagcataagtgtaaagc 1949
QY 1041 -----tgcggtgggctctatgtctct-- 1062
Db 1950 ctgggggtgctaatgtagtgtaactacacattaatgtcggttggtgcctactgcccgttc 2009
QY 1063 -----gagcggaagaagaacagctgcatlaatagaatcgggcaacgcgcgggagagag 1114
Db 2010 ccagctcggaacactgtctgcccagctgcatlaatagaatcgggcaacgcgcgggagagag 2069
QY 1115 cgggttgcatggggcggtctccgcttccgctcactgactgactcggtgtcggtcgt 1174
Db 2070 cgggttgcatggggcggtctctccgcttccgctcactgactgactcggtcgtcgt 2129
QY 1175 tcggtctgcgcgagcggtatcaagctcaactcaaaagcggtatacgttatccacagaatc 1234
Db 2130 tcggtctgcgcgagcggtatcaagctcaactcaaaagcggtatacgttatccacagaatc 2189

QY 1235 aggggataacgcaggaagaacatgtgtagcagaagaagccagcaaaagccaggaacggtaa 1294
Db 2190 aggggataacgcaggaagaacatgtgtagcagaagaagccagcaaaagccaggaacggtaa 2249
QY 1295 aaagccggtgtgtgtgtgttttccatagctccgcgcccccttgacgagcatcacaanaa 1354
Db 2250 aaagccggtgtgtgtgtgttttccatagctccgcgcccccttgacgagcatcacaanaa 2309
QY 1355 tcgacgtcgaagtcaaggtgtgcgaaacccgacagagactaaagataccaagcggttcc 1414
Db 2310 tcgacgtcgaagtcaaggtgtgcgaaacccgacagagactaaagataccaagcggttcc 2369
QY 1415 cccctggagagccctctgtgctctcctcgttccgaccccttgccgcttaccggtatccgttc 1474
Db 2370 cccctggagagccctctgtgctctcctcgttccgaccccttgccgcttaccggtatccgttc 2429
QY 1475 cgccttctcccttcgggaacgctggtccttccatgctcaagctgtaagttactcaag 1534
Db 2430 cgccttctcccttcgggaacgctggtccttccatgctcaagctgtaagttactcaag 2489
QY 1535 ttcggtgtaagtcgttcgctcgaagctgtgtgtgcaagaccccccggttcagccoga 1594
Db 2490 ttcggtgtaagtcgttcgctcgaagctgtgtgtgcaagaccccccggttcagccoga 2549
QY 1595 ccgctgcgcttaccggttaactatcgtcttgaatccaacccggttaagacagacttacc 1654
Db 2550 ccgctgcgcttaccggttaactatcgtcttgaatccaacccggttaagacagacttacc 2609
QY 1655 gccactgcaagcagccactgtgtaacagattagcagagcgaggtatgtagcggtgtac 1714
Db 2610 gccactgcaagcagccactgtgtaacagattagcagagcgaggtatgtagcggtgtac 2669
QY 1715 agagttctgaagtgtgtgcttaactacggtctacactagaagaagcagttatgtgtactg 1774
Db 2670 agagttctgaagtgtgtgcttaactacggtctacactagaagaagcagttatgtgtactg 2729
QY 1775 cgcctgtctgaagcagttactctcggaaaaaaggtgtgtgactccttgatccggcaaaa 1834
Db 2730 cgcctgtctgaagcagttactctcggaaaaaaggtgtgtgactccttgatccggcaaaa 2789
QY 1835 aaccacgcgtgtagcggtgtttttgtgttgtaagcagagattacgcgcagaaaaa 1894
Db 2790 aaccacgcgtgtagcggtgtttttgtgttgtaagcagagattacgcgcagaaaaa 2849
QY 1895 aggatctcaagaagaatccttgaatcttctacgggggtctgaagctcagtgtaagcagaaa 1954
Db 2850 aggatctcaagaagaatccttgaatcttctacgggggtctgaagctcagtgtaagcagaaa 2909
QY 1955 ctcaagttaaaggaattgtgtcatagcgagatataatttgaatgtattaga 2007
Db 2910 ctcaagttaaaggaattgtgtcatagagatatacaaaaaggaatcttcaactaga 2962

RESULT 12
AAZ23778
ID AAZ23778 standard; DNA; 8705 BP.
AC AAZ23778;
XX
DT 14-JAN-2000 (first entry)
XX
DE Vector pShuttle DNA.
XX
KW Antisense; DNA library; identification; multiple cloning site; MCS;
KM inhibition; ss.
XX
OS Synthetic.
XX
PN W0950457-A1.
XX
PD 07-OCT-1999.
XX

CC reproductive and respiratory syndrome (PRRS) virus P129a cDNA (see
CC AA27809) was cloned into the modified vector to create PCMV-S-P129
CC (ATCC 203489), which was used to demonstrate cellular expression of
CC PRRS virus by direct transfection of cDNA into cells. The
CC invention relates to polynucleotide molecules, plasmids, viral
CC vectors and transfectected host cells that comprise North American
CC PRRS DNA. It also relates to polynucleotide molecules, viral
CC vectors and transfectected host cells encoding a genetically modified
CC North American PRRS virus that is disabled in its ability to cause
CC PRRS, or which encodes 1 or more heterologous antigenic epitopes,
CC for use as a vaccine.

XX Sequence 3796 BP; 953 A; 934 C; 956 G; 953 T; 0 other;

Query Match 12.8%; Score 1036.2; DB 21; Length 3796;
Best Local Similarity 71.8%; Pred. No. 1,7e-146;
Matches 1613; Conservative 0; Mismatches 293; Indels 341; Gaps 6;

QY 102 cgttaataacttagcgttaatgcccgcgtgctgtagccgccaagaccccgccatt 161
DB 27 cgttaataacttagcgttaatgcccgcgtgctgtagccgccaagaccccgccatt 86
QY 162 gacgtcataatgacgtatgtcccatagtagcagccaatagggaattccattagcgtca 221
DB 87 gacgtcataatgacgtatgtcccatagtagcagccaatagggaattccattagcgtca 146
QY 222 atggtgtagactattacgtgtaaacctgcccacttgcagcatcatcaagtatatactgccc 281
DB 147 atggtgtagactattacgtgtaaacctgcccacttgcagcatcatcaagtatatactgccc 206
QY 282 aatgtagcccccattagcgtcaatgagtaatgagccgcccgcgattatgcccagta 341
DB 207 aatgtagcccccattagcgtcaatgagtaatgagccgcccgcgattatgcccagta 266
QY 342 caggaaccttagggaattctcactcttgcagcatcatcaatgatactatcgtctatc 401
DB 267 caggaaccttagggaattctcactcttgcagcatcatcaatgatactatcgtctatc 326
QY 402 catgtgtatgcggttttgcagcatcatcaatgagcggtgtagtaagcggtttgaactacgggg 461
DB 327 catgtgtatgcggttttgcagcatcatcaatgagcggtgtagtaagcggtttgaactacgggg 386
QY 462 attccaagctcacaacccattgacgtcaatgaggtttgtttggacacaataacag 521
DB 387 attccaagctcacaacccattgacgtcaatgaggtttgtttggacacaataacag 446
QY 522 ggaattccaanaatgtcgtatacaactcgcgccattgacgcaaatggcggtagcggt 581
DB 447 ggaattccaanaatgtcgtatacaactcgcgccattgacgcaaatggcggtagcggt 506
QY 582 acggtggagagtcataataagcagagctctctggtctaa----- 619
DB 507 acggtggagagtcataataagcagagctctctggtctaa----- 566
QY 620 -----ct 621
DB 567 ccatccaagctgttttgaacctcatagaagacacgggagccgaltccaagctccgagctct 626
QY 622 agagaacccaagctct-----tactgtgcttaacgaatatacagactcatataggaga 676
DB 627 agagaacccaagctctctcgtaggaactgaaacccgaagaattcaactggttaatttagtcttc 686
QY 677 cccaagctgtgtagcagagctcgatccaactagtagaaggccgcaagtgctgtaaatctg 736
DB 687 ttgtcttttatttcagggtccgagctcggtgtgtgtagtaataagaactcgtccag 746
QY 737 cagatatccaacacagcgcggtcggtcgagcatgcatctagaaggccctattctataagt 796
DB 747 tggagtgtgcttactacttagagcctgttagcgaagtgtagcttctgtcttaaaagcgcg 806
QY 797 gtcaacataatgtagagctcgctgtagcagctcgagctgtgctttagttagtgcaagcat 856

DB 807 gaattgaccgcgcgcgccaagatatacgccctagaagaatctcatgtgtaccat 866
QY 857 ctg-----ttgttggccctcccccgtgctctctcttgaacccctgtaagggtgcactcc 908
DB 867 ccgcgaaccccttaattacagcatgagcggaatttaaacagagcccggtactactagtagcg 926
QY 909 cactgctcttcccaataaataagggaattgcatcgcattgctga----- 955
DB 927 gcgcggggatccagacatgataagatacatgtgattggtttggacaaccaactagaa 986
QY 956 ----- 955
DB 987 tgcagtgaanaaatgcttattttgtgaatttgtgtagtgcattatttgaacca 1046
QY 956 -----gtaggtgcattcatcttcgggg 979
DB 1047 ttataagctgcaataaacaagttaacaacaattcatcattatgttttaagttc 1106
QY 980 gtgggtgtggcgagacagcaaggggagagatgtggagacatatagcagcatgctgg- 1038
DB 1107 aggggaggtgtggaaggttttctcgatccctagagtcgaactcgagcatgcaagct 1166
QY 1039 ----- 1038
DB 1167 tggcgtatcatgtgtcatagctgttctctgtgtgaattgtatccgctcaaatccac 1226
QY 1039 ----- 1038
DB 1227 acaacatacagccggaagacataaagtgtaaagcctggtgtcctaaatgtagtgaac 1286
QY 1039 -----gagtcggtgtgctctatagcttctgtagggcggaagaa-----ccagc 1080
DB 1287 tcaacataattgctgtgtgcctcacctgcctccagtcogggaaacctgtgtgcac 1346
QY 1081 tgcataatgtaatcggcacaacgcgcgggagagagcggttttgcattgtggcgctctccg 1140
DB 1347 tgcataatgtaatcggcacaacgcgcgggagagagcggttttgcattgtggcgctctccg 1406
QY 1141 ctctcctgcctcaactgactcgtcgtcgtcgtgtcgtgtcgtgcggagcggtatcagctc 1200
DB 1407 ctctcctgcctcaactgactcgtcgtcgtcgtgtcgtgtcgtgcggagcggtatcagctc 1466
QY 1201 actcaaaagcggttaatacggttatccacagaataagggttaacgacggaaagaactgt 1260
DB 1467 actcaaaagcggttaatacggttatccacagaataagggttaacgacggaaagaactgt 1526
QY 1261 gagcaaaaggccagcaaaaggccaggaacggtlaaaaggcggttgcgtgttttcc 1320
DB 1527 gagcaaaaggccagcaaaaggccaggaacggtlaaaaggcggttgcgtgttttcc 1586
QY 1321 ataggtcgcgcgcgcctgacgagcatcacaanaatcgagctcaagtcaaggttgcgaa 1380
DB 1587 ataggtcgcgcgcgcctgacgagcatcacaanaatcgagctcaagtcaaggttgcgaa 1646
QY 1381 acccggaagacatataaagatacaggcggtttcccccctgtaaggtccctgtgtgcctc 1440
DB 1647 acccggaagacatataaagatacaggcggtttcccccctgtaaggtccctgtgtgcctc 1706
QY 1441 ctgttccgaacctgcgcgttaccggaatacgttgcgccttctcccttcgggaagcggtg 1500
DB 1707 ctgttccgaacctgcgcgttaccggaatacgttgcgccttctcccttcgggaagcggtg 1766
QY 1501 cgccttctcaatgtcgaacgctgtatcctcagttcgggtgtaggtcgttcgcgtcaagc 1560
DB 1767 cgccttctcaatgtcgaacgctgtatcctcagttcgggtgtaggtcgttcgcgtcaagc 1826
QY 1561 tgggtgtgtgtgacgaaccccccgtttagcccggaacggtgcgttcttcggtaaacatc 1620
DB 1827 tgggtgtgtgtgacgaaccccccgtttagcccggaacggtgcgttcttcggtaaacatc 1886
QY 1621 gtcttagtccaaccccgtagaagacagacttatccagctgagcagcagcaactgtgtaca 1680
DB 1887 gtcttagtccaaccccgtagaagacagacttatccagctgagcagcagcaactgtgtaca 1946

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 3, 2001, 15:49:24 : Search time 154.79 Seconds
(without alignments)
9924.800 Million cell updates/sec

Title: US-09-778-516-1

Perfect score: 8115

Sequence: 1 gatgtacgggcagatatatcac.....gcgtgtcgtacgacgagtc 8115

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 324599 seqs, 9465562 residues

Total number of hits satisfying chosen parameters: 649198

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents_NA: *
1: /cgn2_6/prodata/2/lna/5A.COMB.seq: *
2: /cgn2_6/prodata/2/lna/5B.COMB.seq: *
3: /cgn2_6/prodata/2/lna/6A.COMB.seq: *
4: /cgn2_6/prodata/2/lna/6B.COMB.seq: *
5: /cgn2_6/prodata/2/lna/6CTUS.COMB.seq: *
6: /cgn2_6/prodata/2/lna/Backfillseq1.seq: *

* Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1240.4	15.3	3853	3 US-08-801-092-5	Sequence 5, App11
2	1068.6	13.2	4026	3 US-08-801-092-19	Sequence 19, App1
3	1053.6	13.0	4326	4 US-08-760-615-7	Sequence 7, App11
4	981.8	12.1	6253	3 US-08-893-327-15	Sequence 15, App1
5	981.8	12.1	6280	3 US-08-893-327-17	Sequence 17, App1
6	981.8	12.1	6280	3 US-08-893-327-19	Sequence 19, App1
7	973	12.0	4283	1 US-08-343-401A-3	Sequence 3, App11
8	973	12.0	4283	1 US-08-445-265A-1	Sequence 1, App11
9	973	12.0	4283	3 US-08-990-442-1	Sequence 1, App11
10	921.8	11.4	6285	1 US-08-467-420A-49	Sequence 49, App1
11	921.8	11.4	6285	1 US-08-470-110A-49	Sequence 49, App1
12	921.8	11.4	6285	1 US-08-667-769A-49	Sequence 49, App1
13	921.8	11.4	6285	2 US-08-940-371-49	Sequence 49, App1
14	921.8	11.4	6285	3 US-08-637-647-49	Sequence 49, App1
15	921.8	11.4	6285	5 PCT-US95-17082A-49	Sequence 49, App1
16	906.4	11.2	3130	4 US-09-038-141-1	Sequence 1, App11
17	906.4	11.2	3789	4 US-09-075-019-8	Sequence 8, App11
18	906.4	11.2	3918	2 US-08-495-500-2	Sequence 2, App11
19	906.4	11.2	4249	1 US-08-801-092-33	Sequence 33, App1
20	906.4	11.2	4453	1 US-08-770-761A-4	Sequence 4, App11
21	906.4	11.2	4454	2 US-08-738-172-2	Sequence 2, App11
22	906.4	11.2	4525	2 US-08-613-861-2	Sequence 2, App11
23	906.4	11.2	4540	1 US-08-770-761A-6	Sequence 6, App11
24	906.4	11.2	4824	2 US-08-485-139-5	Sequence 5, App11
25	906.4	11.2	4824	3 US-08-750-357-5	Sequence 5, App11
26	906.4	11.2	4883	1 US-08-064-121-4	Sequence 4, App11
27	906.4	11.2	4883	1 US-08-318-772A-1	Sequence 1, App11

28	906.4	11.2	4883	1 US-08-478-015-4	Sequence 4, App11
29	906.4	11.2	4883	3 US-08-475-975-4	Sequence 4, App11
30	906.4	11.2	4883	3 US-09-084-889-4	Sequence 4, App11
31	906.4	11.2	4949	3 US-09-138-024-22	Sequence 22, App1
32	906.4	11.2	5399	1 US-08-064-121-1	Sequence 1, App11
33	906.4	11.2	5399	1 US-08-478-015-1	Sequence 1, App11
34	906.4	11.2	5399	3 US-08-475-975-1	Sequence 1, App11
35	906.4	11.2	5399	3 US-09-084-889-1	Sequence 1, App11
36	906.4	11.2	5618	3 US-08-799-569-1	Sequence 1, App11
37	906.4	11.2	5620	1 US-08-104-072B-9	Sequence 9, App11
38	906.4	11.2	5620	1 US-08-351-413-3	Sequence 3, App11
39	906.4	11.2	5620	2 US-09-025-583-3	Sequence 3, App11
40	906.4	11.2	5642	1 US-08-318-772A-2	Sequence 2, App11
41	906.4	11.2	5653	1 US-08-073-836-3	Sequence 3, App11
42	906.4	11.2	5653	1 US-08-233-277-1	Sequence 1, App11
43	906.4	11.2	5703	1 US-08-467-420A-50	Sequence 50, App1
44	906.4	11.2	5703	1 US-08-470-110A-50	Sequence 50, App1
45	906.4	11.2	5703	1 US-08-667-769A-50	Sequence 50, App1

ALIGNMENTS

RESULT 1
US-08-801-092-5
; Sequence 5, Application US/08801092
; Patent No. 6074850
; GENERAL INFORMATION:
; APPLICANT: Antelman, Douglas
; APPLICANT: Gregory, Richard J.
; APPLICANT: Mills, Kenneth N.
; TITLE OF INVENTION: Tissue Specific Expression of
; TITLE OF INVENTION: Retinoblastoma Protein
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/801,092
; FILING DATE: 14-FEB-1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/751,517
; FILING DATE: 15-NOV-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Flits, Renee A.
; REGISTRATION NUMBER: 35,136
; REFERENCE/DOCKET NUMBER: 016930-001020
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-576-0200
; TELEFAX: 703-576-0300
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3853 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 209..250
; FEATURE:
; NAME/KEY: CDS

```

? LOCATION: 254..289
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 293..505
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 509..514
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 518..520
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 524..658
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 662..691
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 695..748
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 752..781
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 785..829
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 1132..1134
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 1138..1149
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 833..862
US-08-801-092-5

```

Query Match	15.3%	Score 1240.4;	DB 3;	Length 3853;
Best Local Similarity	74.1%;	Pred. No. 1e+68;		
Matches 1911; Conservative	0;	Mismatches 96;	Indels 573;	Gaps 4;

[illegible]

QY	481	atcgaagcgaatgaggagttgttllttgycaccacaatacgaaggaacttccaaatgctgt	540
DB	690	attgacgcgcaatggagactttgttttggcaccacaaaatcaacgggagactttccaaaatgtcgt	749
QY	541	aagaccccgcccaattggaagcaaatggggcggtgaagcggtgacggtgggaagttacata	600
DB	750	aacacactccggcccatgagcgaataaggcggtgagcggtgacggtgggaagttacata	809
QY	601	agcagagctctctgcttaactagaagaccacactgcttactggtcttatcgaaattaaacg	660
DB	810	agcagagctctctgcttaactagaagaccacactgcttactggtcttatcgaaattaaacg	869
QY	661	actcaataaggagagaccacaagctt-----	685
DB	870	actcactataggagagaccacaagctt-----	929
QY	686	-----	685
DB	930	cgagggccagctgttggggcgcggttgagacaaacttccgcggtctttccagctac	989
QY	686	-----gtaacgagctcgatccactagtaaacgcccgaagtgc	726
DB	990	ttggatcggaaccccgccgcgctccgacagcgtactccgaccagaggaactcgacagcgt	1049
QY	727	tgaaattcgcagatalcatalcactgagcgcg-----tcgag	767
DB	1050	ccgacatcgacggatcggaamaacctctcgagggcgccgctgcagctagacgaattccgcg	1109
QY	768	catgacatctagagggccctattctatagtgtaacctaaatgctagaagctcgtagcag	827
DB	1110	tacgatatcgatggggccctattctatagtgtaacctaaatgctagaagctcgtagcagc	1169
QY	828	ctcgactgctgctcttagtttgcagccacatctgtgtgttgcgcctcccgctgctctct	887
DB	1170	ctcgactgctgctcttagtttgcagccacatctgtgtgttgcgcctcccgctgctctct	1229
QY	888	gaccctggaagtgccacatcccaactgctcttccctaaataaaatgagaaattgacgca	947
DB	1230	gaccttggaagtgccacatcccaactgctcttccctaaataaaatgagaaattgacgcca	1289
QY	948	tttgtctgagtgatggtcatcttctatcttcgggggggtgggtgagggcaggaagaaggga	1007
DB	1290	tttgtctgagtgatggtcatcttctatcttcgggggggtgggtgagggcaggaagaaggga	1349
QY	1008	ggattgggaagacatatagcagcatgcttgggga-----	1040
DB	1350	ggattgggaagacatatagcagcatgcttgggga-----	1409
QY	1041	-----	1040
DB	1410	attttcgattccaccgcccgcctctatgaaagggttgggcttgggaattggtttccggagc	1469
QY	1041	-----	1040
DB	1470	ccggctgagatgatcctccacggcgggagatctatgctgagattcttccgccaccaccaact	1529
QY	1041	-----	1040
DB	1530	tgttattatgcagcttatataatggtttacaaataaagcaatagcatcacaaatttcacaata	1589
QY	1041	-----	1040
DB	1650	atgtctgatatccctgcagactcttagctagagcttgscgttaatcatagctatagctgttc	1709
QY	1041	-----	1040
DB	1710	ctgtgtgaaattgtttatccgctacaaattccacacaaactagagacggcgaacacataaagt	1769
QY	1041	-----tgcggtgggctctatg	1057

Db	1770	GTAAAGCCTGGGGTGCTTAATGATGAGCTAACTACATTAATTCTCGTTGGCTACTACG				1829
Qy	1058	ctctc-----*gagcggaagaagaccagctgcatltaatgaaicgcgcaacgcyg				1107
Db	1830	CCGGTTTTCAGTGGGAAACCTGTCGTGGCAGGTGATTAATGATTCGGCAAGCGCGG				1889
Qy	1108	ggaagagcggttgcgtatctgggcgtctctccgctctcctgcctcaactgcgtgcgt				1157
Db	1890	GGAAAGCGGCTTTTTCGTAATTTGGGCGCTTCCTCCGCTCCTCACTACACTGCGTGCCT				1949
Qy	1158	cggtcgtcttcgagctcgagcggtatcaagctcaactcaaaagcggtaaaggtatcca				1227
Db	1950	CGGTCTTCGGCTTCGGCGAGCGGTTATCAGCTCCTCAAAAGCGGTAATACGGTTATCA				2009
Qy	1228	cagaaatcagggataacgcagaaagaacatgtgacaaagggccaganaaaggccagaa				1287
Db	2010	CAGATCAGGGGATACGACGCAAAACATGTGAGCAAAAGCCAGCAAAAGGCCAGGA				2059
Qy	1288	accgtlaaaagcgcgcttgcgtcttcttcataaggtctcgcccccttaagagatc				1347
Db	2070	ACCGTAAAGCCGCGCTTCTGCGCTTTTTCATRAGCTCGCCGCCCTTGACGACATC				2129
Qy	1348	acaaanaicgagcgtcaagtcagaggttggcgaaccgcagagactataaagataccag				1407
Db	2130	ACAAAAATCGACGCTCAATGACAGAGGTGGGAAACCCGACAGACTATTAAGATACAGG				2189
Qy	1408	cggttcccccttgagaagctccctcgtgcgtctcctggtttccgaacctgcgcttacggat				1467
Db	2190	CGTTTCCCCCGGAAGTCCCTGTCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT				2249
Qy	1468	accgttcgccttcttcccttcggaagcgctgagcgcttcttcaatgctcaacgctgtagt				1527
Db	2250	ACGTGTCGCGCTTCT				2309
Qy	1528	atctcagatcggtgtagtgcgttcgtctccaagctggcgctgctgacgaaccocccgctc				1587
Db	2310	ATTCATATTTGGGTGTAGTGCTTCCGCTCCAAAGCTGGGCGTGTGTGACCAACCCCGCTTC				2369
Qy	1588	agcccgagcgctgcgacctctccgtaactatcgcttbgagtcgaaccocgctlaagacg				1647
Db	2370	AGCCCGAACCGCTGCGCTTATCCGGTACTATCGTCTTGATGCCAACCCGGTAACACG				2429
Qy	1648	actatcgcacatgagcagcagccactgtgtaagaagattagcagagcgagatgtatgagcg				1707
Db	2430	ACTTATGCGCACTGGCAGCGACCACTGTTAACAGGATTAGCAGACGAGATGATGAGCGC				2489
Qy	1708	gtgtactaagaagttcttgaagtggtggtcctaactaagcgtcaactagaagagacgatttg				1767
Db	2490	GTGTACTAAGAGTTCTTGAATGTGTGCTTACTACGCGCTACCTGTGAAGAGACAGATTG				2549
Qy	1768	gtactcgtgcctcgtcgtgaagcagcgttacccttcggaanaagattgtagtccctgatacg				1827
Db	2550	GTATCTCGCGCTGTGCTGAACCGATTAACCTTGGAAAAAGATTGTTGATCTTGATCCG				2609
Qy	1828	gcaaaacaaaccacgctgtgtagcggtgtgttttttggtttgcaagcagacgattacogca				1887
Db	2610	GCAAAACAAACACCGCTGTGATGGGCTTTTGTGTTTGTGCAAGCAGCAGATTACGCGCA				2669
Qy	1888	gaaaaaagagatctcaagaagatccttltgataccttctcaacgaggtcgtgaagctgaatga				1947
Db	2670	GAAAAAAGGATCTCAAGAAGATCTTTGATCTTTTCTTACGGGGCTTACCGCTCAGTGA				2729
Qy	1948	acgaanaaccagttlaaggaatttggtcagaagcagagabacatalttgaatglattaga				2007
Db	2730	ACGAAAACTACGTTAAGGATTTTGTATGATGATATTACAAAAAGATCTTCACTACATTTGA				2789

RESULT 2
US-08-801-092-19
; Sequence 19, Application US/08801092
; Patent No. 6074850
; GENERAL INFORMATION:

```

1  APPLICANT: Antelman, Douglas
2  APPLICANT: Gregory, Richard J.
3  APPLICANT: Wills, Kenneth N.
4  TITLE OF INVENTION: Tissue Specific Expression of
5  TITLE OF INVENTION: Retinoblastoma Protein
6  NUMBER OF SEQUENCES: 46
7  CORRESPONDENCE ADDRESS:
8  ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP
9  STREET: Two Embarcadero Center, 8th Floor
10 CITY: San Francisco
11 STATE: CA
12 COUNTRY: USA
13 ZIP: 94111
14 COMPUTER READABLE FORM:
15 MEDIUM TYPE: Floppy disk
16 COMPUTER: IBM PC compatible
17 OPERATING SYSTEM: PC-DOS/MS-DOS
18 SOFTWARE: Patent Release #1.0, Version #1.30
19 CURRENT APPLICATION DATA:
20 APPLICATION NUMBER: US/08/801,092
21 FILING DATE: 14-FEB-1997
22 CLASSIFICATION: 514
23 PRIOR APPLICATION DATA:
24 APPLICATION NUMBER: US 08/751,517
25 FILING DATE: 15-NOV-1996
26 CLASSIFICATION: 514
27 ATTORNEY/AGENT INFORMATION:
28 NAME: Fitts, Renee A.
29 REGISTRATION NUMBER: 35,136
30 REFERENCE/DOCKET NUMBER: 016930-001020
31 TELECOMMUNICATION INFORMATION:
32 TELEPHONE: 415-576-0200
33 TELEFAX: 703-576-0300
34 INFORMATION FOR SEQ. ID NO: 19:
35 SEQUENCE CHARACTERISTICS:
36 LENGTH: 4026 base pairs
37 TYPE: nucleic acid
38 STRANDEDNESS: single
39 TOPOLOGY: linear
40 MOLECULE TYPE: DNA (genomic)
41 FEATURE:
42 NAME/KEY: CDS
43 LOCATION: 209..250
44 FEATURE:
45 NAME/KEY: CDS
46 LOCATION: 254..289
47 FEATURE:
48 NAME/KEY: CDS
49 LOCATION: 293..505
50 FEATURE:
51 NAME/KEY: CDS
52 LOCATION: 509..514
53 FEATURE:
54 NAME/KEY: CDS
55 LOCATION: 518..520
56 FEATURE:
57 NAME/KEY: CDS
58 LOCATION: 524..658
59 FEATURE:
60 NAME/KEY: CDS
61 LOCATION: 662..691
62 FEATURE:
63 NAME/KEY: CDS
64 LOCATION: 695..748
65 FEATURE:
66 NAME/KEY: CDS
67 LOCATION: 752..781
68 FEATURE:
69 NAME/KEY: CDS
70 LOCATION: 785..829
71 FEATURE:
72 NAME/KEY: CDS
73 LOCATION: 833..862

```

FEATURE: CDS
NAME/KEY: 1305..1307
LOCATION: 1305..1307
FEATURE: CDS
NAME/KEY: 1311..1322
LOCATION: 1311..1322
US-08-801-092-19

Query Match 13.2%; Score 1068.6; DB 3; Length 4026;
Best Local Similarity 69.7%; Pred. No. 3.2e-230;
Matches 1918; Conservative 0; Mismatches 89; Indels 746; Gaps 5;

QY 1 gatgtacggcgacagatatacagcgttgaacatgtatattgactagttatataagtaacatca 60
DB 210 GATGTACGGCGACAGATATACCGGTTGACATTGATTAGTATTATATAGTAATCA 269
QY 61 attacggggtcattagttacatagccacataatgaagtcgcggttaacataactaggtta 120
DB 270 ATTACGGGGTCTATTAGTTTCATAGCCCATATATAGAGTTCCCGTTACATTAACGTTA 329
QY 121 aatggccgcctgtgctgaacgcccacgaaccccgccattgaagtcacataatgaagctat 180
DB 330 AATGGCCCGCTGTGCTGACCCGCCACAGACCCCGCCCATTTGACGTCAATATGACGTAT 389
QY 181 gtcccatagtaacagccaatagggactttccattgaacgtcaatgggtgagactattacg 240
DB 390 GTCCCATAGTAACGCCCAATAGGAGCTTTCATTGACGTCAATGAGGTGAGACTATTTACGG 449
QY 241 taactgcgcacactgtgacgtatcatcaatgatatataatgccaaagtaagcccccattac 300
DB 450 TAACTGCCCACTGTGAGCTGATCAATCAAGTTATATGATGCCAAGTACCCCTTATTTAC 509
QY 301 gtcaatgacggttaaatgacgcccgcctgcatlataagccagtaacatgactatgagactt 360
DB 510 GTCAATGACGGTTAAATGAGCCGCCCTGSCATTATGCCCATATACCTTATGAGCACTT 569
QY 361 cctactgtgacgtacatctacgtatagtcacgtatcgtatcattacatggtgaggtttgg 420
DB 570 CCTACTGTGACGTACATCTACGTATTAGTACGCTATTACCATGATGATGGGTTTGG 629
QY 421 cagtaacatcaatgagcggtgagtagcggttgactacagggagattccaaagtcacaccc 480
DB 630 CAGTACATCAATAGGCGGTGATAGCGGTTTACACGCGGATTTCCAAAGTCTCCACCC 689
QY 481 attgacgtcaatgaggtgtgtttgtgaccaaataacaaacagggacttccaaaagtgct 540
DB 690 ATTGACGTCAATGAGAGTTGTTTGTGGACCAAAATCAACGAGACTTCCAAAAGTGTCT 749
QY 541 aacaactcgcgcccaatgagcaaaatgggagtagcggtgaggtgaggtgaggttataata 600
DB 750 AACAACTCCGCCCATTTACGCAAAATGGCGTAGCGCTGTACGGTGGAGGTCTATATA 809
QY 601 agcagagatcctgtgctacactagaacacacactgtctactgtctatgaaatataacg 660
DB 810 AGCAGAGATCTCTGTGGCTACTAGAGAACCCACTGCTTACTGGCTTATGAAATTAATACG 869
QY 661 actactactaagaggaaccaagcttg----- 687
DB 870 ACTGACTTATAGGAGAACCAAGCTTCGCGCGGTACCACTCTTCCGCAATGCTGTCTG 929
QY 688 ----- 687
DB 930 CGAGGGCCAGCTGTGGGCTCGCGGTGAGGACAACCTTCGCGGTTCCTCCAGTACTC 989
QY 688 ----- 687
DB 990 TTGGATGGAAACCCGTCGCCCTCCGAACGCTACTCGCCACCGAGGAGCCTGAGCGAGT 1049
QY 688 ----- 687
DB 1050 CCGCATGACCGGATCGGAAAACTCTCGAGGAGACTGAAAAACGAAAGTTAACTGGTA 1109

QY 688 -----taacgagctggatccactagtaacgagccgcaatg 723
DB 1110 AGTTAGTCTTTTGTCTTTTATTTATTCAGTCCCGCATCCGCTGTGTCATAATCAAG 1169
QY 724 tgcgtgaattctgcagatat----- 743
DB 1170 AACTGCTCTCAAGTGGATGTGGCTTTACTTCTAGGCTGTACGAACTGTTACTTCTGC 1229
QY 744 -----ccatcacactggcgagccgc-----tcgagcatgcat 774
DB 1230 TCTAAAGCTGGGAAATTTGTAACCGCGCGCGGTGACGTCFAGAGCAATTCGGTACGATA 1289
QY 775 ctagaaggccctatctatagtgtaacctaaatgctaagctcgtgtatcaagctcgtact 834
DB 1290 TCGATGCGCCCTATCTATGATGTACCTAAATGCTATAGAGTCGCTGATCGCTCGACT 1349
QY 835 gtgcctctagttgcgaacacatgctgtgtgtggccctcccgctgacctcttgaacccgt 894
DB 1350 GTGCCTTATAGTGGCACCAATCTGTGTGTGGCCCTCCCGGTGCTTCCCTTGACCCCTG 1409
QY 895 gaaagtgcacctcccaactgctccttcccaataaataaggaatgcatcgcattgtctg 954
DB 1410 GAAGGTGCACTCCCACTGCTCCTTCTTAATTAATGAGAAATTCATCGCATTTGCTGTG 1469
QY 955 agtaagtgcatctatctctgtgggggtggtggtggtggtggtggtggtggtggtggtggt 1014
DB 1470 AGTAGGTGTCAATTCATTCGTGGGGGTGGGGGTGGGGGTGGGGGTGGGGGTGGGGGTGG 1529
QY 1015 gaaagcaataagcaagcagtcgctgggga----- 1040
DB 1530 GAAGCAATAGCCGAATGACCGACCAAGCGACGCCCAACTGCCATCAGAGATTTCGA 1589
QY 1041 ----- 1040
DB 1590 TTCCACCGCGCTCTTAATGAAGGTGGCTTCGGAATCGTTTCCGGAGCGCGCTG 1649
QY 1041 ----- 1040
DB 1650 GATGATTCCTCAGCGCGGGGATCTCATGCTGAGTTCTTCCGCCCACTTGTTTAT 1709
QY 1041 ----- 1040
DB 1710 TGACGCTTATATGTTGTACAAATAAGCAATAGCATTCACAAATTTACAAATAAGCAT 1769
QY 1041 ----- 1040
DB 1770 TTTTCACTGCATCTAGTGTGTTGTGCCAACTCATCATATGATCTTATCATGTCTG 1829
QY 1041 ----- 1040
DB 1830 TATACCGTCGACCTCTAGCTAGAGCTTGGCTTAATCATGTCATAGCTTCTGTGTG 1889
QY 1041 ----- 1040
DB 1890 AAATTGTTATCCGCTACAAATTTCCACACACATACGAGCCGGAACATTAAGTGAAGC 1949
QY 1041 -----tgcggttgggtcctatggtctt-- 1062
DB 1950 CTGGGTCCTTAATGATGAGCTAATCTACATTAATTCGTTGCGCTCAGTCCCGCTTT 2009
QY 1063 -----gaagcggaagaagacagctgcatatgaatcggtgcaagcggtgggaaggg 1114
DB 2010 CCAGTCGGGAAACCTGTGCTGCACCTCATTAATTAAGTACGCGCAACGCGGGGAGAG 2069
QY 1115 cggttgcgtatctgggagctcttcgcgtcctcgctacactgactgcgtgcgtgcgtgcgt 1174
DB 2070 CGGTTGCGTATTTGGGCGCTCTTCGCGCTTCTCGCTCACTGACTGCTGCGCTGCGTGT 2129
QY 1175 tgcgttcggtgcgagcgtatcagctcactcaaaagcggtlaatacgttataccacagaaatc 1234
DB 2130 TCGGCTGGCGGAGCGGATACGATCACTCAAAAGCGGTTAATACGTTATCCACGAATTC 2189
QY 1235 aggggatacgcaggaagaacatgtaggcaaaaagccagcaaaaagccaggaacggttaa 1294

|||||
Db 2190 AGGGGATTAACGACGAGAAAGACATGTAGCAAAAAGCCACAAAAGCCGACGAAACCTTAA 2249
QY 1295 aaagccgctgtctgtgctgttccataagctccgccccctgacgagcatcaaaaa 1354
|||||
Db 2250 AAAGGCGCGTCTGCGGTTTTCATAGGCTCCGCCCTGACGAGATCACAAAA 2309
QY 1355 tcgagctcaagtcagaggtgagcgaaacccgacagagctataagatacaaggcgtttcc 1414
|||||
Db 2310 TCGACGCTCAAGTCAGAGGTGGGAAACCCGACAGACTATAAGATACAGGCGTTTCC 2369
QY 1415 cccgtgaagctccctgtgctctctctgttccgaacctccgcttaccgagatacgttc 1474
|||||
Db 2370 CCTGGAAGCTCCCTCGTGGCTCTCTGTTCGACACCTCCGCTTACCGGATACCTGTC 2429
QY 1475 cgccttctcccttcggaaagctgtgacgttctcctaagtctcagctgtatgatacag 1534
|||||
Db 2430 CGCCTTCTCCCTTGGGAGCGTGGCGCTTCTCATGTGCTCAAGCTGTAGTATCTGAG 2489
QY 1535 ttggtgtgtagtgcgttcgctcgaagctggtgtgtgtgcaagaacccccgttcaagccga 1594
|||||
Db 2490 TTGCGGTGATAGTGTGCTGCTCAAGCTGGGCTGTGTGACAGAACCCCGTTCAGCCGA 2549
QY 1595 cgcgtgaccttaccggtactacgtctgttgaatcccaacccgtaagacagactatc 1654
|||||
Db 2550 CCGCTGGCGCTTATCCGTTAACTATCTGTGAGTCCCAACCCGTTAGACACGACTTATC 2609
QY 1655 gccactgagcagcagccactgttacaagattagcagagcagagtlatgagcggtgtac 1714
|||||
Db 2610 GCCACTGGCAGCAGCCACTGTAAACAGATTAGACAGCAGGATATAGGCGGTGCTAC 2669
QY 1715 agagttcttaagtgtgtgacctactacgtctacagagaagaatattgttactctg 1774
|||||
Db 2670 AGAGTTCTTAAAGTGAGGCTTAACAGGCTACACTAGAGGACAGTATTGTGATCTG 2729
QY 1775 cgcctgtctgaagcagcttacccttcggaagaagttgtgtagcttctgtaccgcaaca 1834
|||||
Db 2730 CGCTCTGCTCAACGCACTTACCTTCGAAAAAGATTGGTAGCTTGTGATCCGCAACA 2789
QY 1835 aaccacgcgtgtgtagcggtgttttttctgttgcgaagcagagattacgcgcagaaaa 1894
|||||
Db 2790 AACCAACCGCTGTAGCGGTGTGTTTGTGTTGCAACGACAGATTACGGCGCAGAAAA 2849
QY 1895 agagatccaagaagatcccttgcattcttctacgggtgtcgaagctagtggaagaaaa 1954
|||||
Db 2850 AGGATCTCAAGAAAGATCCTTGTATCTTCTACGGGCTGTGACGCTCAGTGAACGAAA 2909
QY 1955 ctacagttgaagattgtgcatgacgagatacatattgaatgatataga 2007
|||||
Db 2910 CTCACGTTAAAGGATTTTGGTTCATGAGATTATCAAAAAGATCTTCCACTAGA 2962

RESULT 3
US-08-760-615-7
Sequence 7, Application US/08760615
Patent No. 6200959
GENERAL INFORMATION:
APPLICANT: Haynes, Joel R
APPLICANT: Schmaljohn, Connie S
APPLICANT: Fuller, Deborah L
APPLICANT: Schmaljohn, Alan
APPLICANT: Jahrling, Peter B
TITLE OF INVENTION: GENETIC INDUCTION OF ANTI-VIRAL IMMUNE
TITLE OF INVENTION: RESPONSE AND GENETIC VACCINE FOR FILOVIRUS
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Quarles & Brady
STREET: 1 South Pinckney Street
CITY: Madison
STATE: WI
COUNTRY: US
ZIP: 53703
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/760,615
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Berson, Bennett J
REGISTRATION NUMBER: 37094
REFERENCE/DOCKET NUMBER: 110229, 91241
TELECOMMUNICATION INFORMATION:
TELEPHONE: 608-251-5000
TELEFAX: 608-251-9166
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 4326 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "Expression vector"
IMMEDIATE SOURCE:
CLONE: pMRG7077
FEATURE:
NAME/KEY: promoter
LOCATION: 1250..2062
FEATURE:
NAME/KEY: Intron
LOCATION: 2063..2887
OTHER INFORMATION: /function= "Human Cytomegalovirus"
FEATURE:
NAME/KEY: polyA-site
LOCATION: 2912..3314
FEATURE:
NAME/KEY: CDS
LOCATION: complement (299..1114)
US-08-760-615-7

Query Match 13.0%; Score 1053.6; DB 4; Length 4326;
Best Local Similarity 94.1%; Pred. No. 7, 6e-227;
Matches 1118; Conservative 0; Mismatches 39; Indels 31; Gaps 1;

QY 820 tgatcaagcctcgactgtgcttctgaattgccaagccatctgttggccctcccccgtg 879
|||||
Db 3030 TGATCAGCCTCGACTGTGCTTCTAGTTGCCAGCCATCTGTTGTTCCCTCCCGCTG 3089
QY 880 ccttcttgaccttggaaggtgccaactccactgttcccttccataaaaaatgaggaatt 939
|||||
Db 3090 CCTTCCTTGAACCTGGAAGGTGCCACTGCCACTGCTCTTCTTAATAAATGAGGAATT 3149
QY 940 gcatcgcatgtctgaatgagtgatcattctctctgtgggtgtgggtgaggaagc 999
|||||
Db 3150 GCATCGCATTTGTGAGTGAAGTGCATCTTCTTGAGGGGTGGGTGGGACAGC 3209
QY 1000 aaggggaagatttggaagaacaatagcagcagcatgctgggagtgagtgctatgct 1059
|||||
Db 3210 AAGGGGAGATTGGGAAGACAATAGCAGCATCTGGGATGGGTGGGCTCATGGCT 3269
QY 1060 tctgaagcggaaagaaacacagctgcatatgaatcggccaacgctgggagagcggtt 1119
|||||
Db 3270 TCTGAGCGGGAAGAACACAGCT 3298
QY 1120 tgcgtatggcgctcttcgcgttccctgcgtcaactgactgcgtgagctggtgtgtggc 1179
|||||
Db 3299 GACAGCTCGACTCTGAATGCTTCCCTGCTCAGCTGACTGCTGCTGCTGCTGCTGGC 3358
QY 1180 tgcggagcaggtatcagctcaactcaaaagcgggtaaatcaggttaacagaatcaagg 1239
|||||

```
Db 3359 TCGGCGAGACGGTATCAGCTCACTCACTCAAAAGCGGTATATACGGTTATCCACAGATCAGGGG 3418
Oy 1240 ataagcaggaagaaacacatgttagcaaaaagccagcaaaaagccaggaacccgttaaaag 1299-
Db 3419 ATAAGCGAGAAAGAAACATGTAGCAAAAAGGCCAGCAAAAAGCCAGAAACCTTAATAAGG 3478
Oy 1300 ccgcgtgctgagcgttttccalaagctccgccccctgacagacatcaaaaatcgac 1359
Db 3479 CCGCGTGTGCTGGCTTTTTCATAGGCTCCGCCCTGACGAGATCCAAATAATCGAC 3538
Oy 1360 gctcaagtcagaggtggtgggaaaccccgacagacataaaagacagagcgtttccccc 1419
Db 3539 GCTCAAGTCAAGGTGGGCAAAACCCGACAGGACTATAAAGATACAGGGGTTTCCCGCTG 3598
Oy 1420 gaagtcctctgctgctctcctgcttccagacctccgcttaccagatccgttccgct 1479
Db 3599 GAAGTCTCCTCTGCTGGCTCTCTCTGTTCCAGACCTGCGCTTACCGGATACCTGCTCGCT 3658
Oy 1480 ttccctctcggaagcgtgagcgttctcctcaatgctcaagcgttagtatactcagttcgg 1539
Db 3659 TTCTCCCTTCGGGAAGCGTGGCGCTTTCATAGCTCACGCTGTAGGTATCTCAGTTCCG 3718
Oy 1540 tgaagtcgttccgtcccaagcgtggtctgtgtgacgaaccccccgcttccagccgct 1599
Db 3719 TGTAGTGTGCTGCTCAAGCTGGGCTGTGTGACGAAACCCCGTTAGCCCGACCGCT 3778
Oy 1600 gcgccttaccgtaactacgtcttgaagtcacacccggttaagacacagacttaccgac 1659
Db 3779 GCGCTTATCCGCTAAGCTATCTCTTGAAGTCCAAACCGGTAAAGACAGACTTATGCGCC 3838
Oy 1660 tggcagcagcactggttaacaggaattagcagagcagagtatgtagcgtgtctacagagt 1719
Db 3839 TGGCAGCAGCCACTGTAAACAGAGATTAGCAGCAGCAGGATATAGCGGCTGTACAGAGT 3898
Oy 1720 tcttaagttgtggtcctacacacagccttcaaggaagaacattttagtctcgtc 1779
Db 3899 TCTTAAGGAGGGGCTTACTAGCTACAGCTACACTAGAAAGAACAGTATTTGGTATCTCGGCTC 3958
Oy 1780 tgcctaaacacacttacccttcgaaaagagttgtagctctgtatccggcaacaaca 1839
Db 3959 TGCTAAACCACTTACCTTTCGAAAAAGTGTAGCTTGTATCCGGCAACAACA 4018
Oy 1840 ccgcgtgtagcgtgtgttttttlttgcaagcagacattacgcgcagaaaaaagat 1899
Db 4019 CCGCGTAGAGCGGTGTTTGTGTCAGACAGCAGATTACGGCGCAAAAAAGCAT 4078
Oy 1900 ctcaagaagatccttgaatcttcttaaggggtcgaagcgaagtggaagcaaaactaac 1959
Db 4079 CTCAGAAAGATCTTGTGATCTTCTACGGGCTGACGCTAGTGGAACGAAACTCAC 4138
Oy 1960 gtaaggaattgtgcatgagcagatacatattgaatgattaga 2007
Db 4139 GTTAAGGATTTTGTGTCATGATATCAAAAAAGATCTTCACCTAGA 4186
```

RESULT 4
US-08-893-327-15

Sequence 15, Application US/08893327

Patent No. 6020192

GENERAL INFORMATION:

APPLICANT: Zolotukhin, Sergei

APPLICANT: Hauswirth, William W.

APPLICANT: Muzyczka, Nicholas

TITLE OF INVENTION: Humanized Green Fluorescent Protein

TITLE OF INVENTION: Genes and Methods

NUMBER OF SEQUENCES: 20

CORRESPONDENCE ADDRESS:

ADDRESSEE: Arnold, White & Durkee

STREET: P. O. Box 4433

CITY: Houston

STATE: TX

COUNTRY: USA

ZIP: 77210-4433

```
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/893,327
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/588,201
FILING DATE: 18-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Kitchell, Barbara S.
REGISTRATION NUMBER: 33,928
REFERENCE/DOCKET NUMBER: UFLA.062\KIT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (713) 789-2679
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 6253 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 988..1701
US-08-893-327-15
```

Query Match 12.1%; Score 981.8; DB 3; Length 6253;
Best Local Similarity 86.9%; Pred. No. 1.1e-210;
Matches 1162; Conservative 0; Mismatches 37; Indels 138; Gaps 2;

```
Oy 809 ctaagagctcgatcagcctcagcttgcctttagttagtgcagccatctgttctgc 868
Db 3017 CTAGAGCTCGCTGATCAGCTCGACTGCTTCTAGTTGCGAGCAATCTGTTTGGC 3076
Oy 869 cctcccccgtgctctcttgcacccctggaaggtgcaaccctccttccctaaataa 928
Db 3077 CTTCCCGGTGCTTCTTGTGACCTGTGAAGGTGCCACTGCCACTCTCTTCTATATAA 3136
Oy 929 atgagaaattgcatcgcattgtcgaagtagtgcattctatctctgggggtggtggtg 988
Db 3137 ATGAGGAATTCATTCGATGTGTGTGAGTAGTGTGATTTCTTGTGGGGGTGGGGG 3196
Oy 988 ggcaggaacagcagggggaagatttggaagacataagcagcagtctggga----- 1040
Db 3197 GGCAGGACAGCAAGGGGGGAGATTGGAAACATAGCAGCATCTGGGAGAGATCTA 3256
Oy 1041 ----- 1040
Db 3257 GGAACCCCTAGTGAATGAGTTGGCCACTCCCTCTGCGCGCTCGCTGACTGAGGCG 3316
Oy 1041 -----tgcygtgggtctatgtctcttgaggcggaaagaccag-- 1079
Db 3317 GCGCGGGGCAAGCCCGGGGCGTGGGGCGACCTTGTGGCCCGGCTCAGTGAGGAGCGG 3376
Oy 1080 -----ctgataatg 1090
Db 3377 AGCGGCGAGAGAGGAGTGCCCAACCCCGCCCGCCCGCTGACGCGCTGACATTAATG 3436
Oy 1091 aatggcaaacgagggggaaggggttggtaattggggcgtcttcgcttccgtcgt 1150
Db 3437 AATCGGCAACCGCGGGGAGAGCGGTTTGGAATTTGGGCGCTTCCGCTCTCCCT 3496
Oy 1151 cactgactcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 1210
Db 3497 CACTGACTCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3556
Oy 1211 gtaatacgttatcacagaatcagggataacgcaagaaacacatgtgaacaaag 1270
```

Db 3557 GGTAATACGGTATTCACAGATCAGGGGATTAACGACGAAAGACATGTAGCAAAAGG 3616
QY 1271 ccaagcaaaagccaggaaccgttaaaagccggttgcgtgttttccatagctccg 1330
Db 3617 CAGCAAAAGCCAGGAAACGTAAGGCGCGGTGCTGGCTTTTTCATAGGCTCCG 3676
QY 1331 cccccctgaagcatcacaaaaatcgaagctcaagtcagagtgagcaaccgagcag 1390
Db 3677 CCCCCCTGAGAGCATACACAAAATGACGCTCAAGTCAAGGTGGGAAACCCGACAGG 3736
QY 1391 actataagaatcacagcgcttcccccctgaagctccctcgtgcgtctcgttccgac 1450
Db 3737 ACTATAAGATACAGGCGCTTCCCTCGAAGCTCCCTCGGCGTCTCCTGTTCGAC 3796
QY 1451 ccgcgcgctaacggtatccctgtccgttccctccttggaagcgtagcgcttctca 1510
Db 3797 CCGCGCGCTTACCGGATCCTGTCCGCTTTCCTTCCGGAAGCGTGGCGCTTCTCA 3856
QY 1511 atgctaacgctgaatcatcctcgtgtgtagtgcgttcctcaagctggagctgt 1570
Db 3857 ATGCTCAGCTGTAGGATATCTCAGTTGCGTGTAGTGTTCCTCCAGCTGGGCTGT 3916
QY 1571 gcaagaacccccctgaagccgacgctgcgtccttaccgtaactatcgtcttgaatc 1630
Db 3917 GCACGAACCCCCGCTTACGCGCCGACCGCTGCTTATCCGCTAATGCTGTGAGTC 3976
QY 1631 caaccggtgaagcaagctatcgcgcaactgcgcaagcgaactggttaagaattagcag 1690
Db 3977 CAACCCGGTAAGACAGACTTATCGCACAGCAGCCACTGTAACGATTAAGCAG 4036
QY 1691 agcagagatgtaggcgtgtcctacagatctctgaagtgtagccttaactacagcgctaac 1750
Db 4037 AGCGAGGATGTAGCGGCGTCTCAGAGTTCTTGAAGTGTGCGCTTACACTACGCTACAC 4096
QY 1751 tagaaggaacgatctgtgtatcgtcgtcgtcgtgaagccagcttaccctcggaaaaagagt 1810
Db 4097 TAGAAGGACAGTATTTGGTATCTGCGCTGCTGTGAAGCCAGTTACCTTCGGAAGAGAGT 4156
QY 1811 tggtagctctgtgctcgcgaacaaacacacgctgtgtagcggtgttttttttgcaaa 1870
Db 4157 TGGTACTCTGTGATCCGCGCAACACACCGGTGTAAGCGGTGTTTGTGCAAA 4216
QY 1871 gcaagcagatcagcgacagaaaaaagatcctcaagaagatccttcttctcagcg 1930
Db 4217 GCAGGAGATTACCGCGCAAGAAAAAGATCTCAGAAAGATCTTTGATCTTTCTACGGG 4276
QY 1931 gctcgaagcgtcagtggaacgaacactcaagcttaaggatcttgcataagcgagatcat 1990
Db 4277 GTCGTGACGCTCAAGTGGAACGAAACTCAGGTAAAGGATTTTGGTATGAGATTATCAAA 4336
QY 1991 attgaatgtatttaga 2007
Db 4337 AAGGATCTTCACCTAGA 4353

RESULT 5
US-08-893-327-17
Sequence 17, Application US/08893327
Patent No. 6020192

GENERAL INFORMATION:
APPLICANT: Zolotukhin, Sergei
APPLICANT: Hauswirth, William W.
APPLICANT: Wuzyczka, Nicholas
TITLE OF INVENTION: Humanized Green Fluorescent Protein
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P. O. Box 4433
CITY: Houston
STATE: TX
COUNTRY: USA
ZIP: 77210-4433

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/893,327
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/588,201
FILING DATE: 18-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Kitchell, Barbara S.
REGISTRATION NUMBER: 33,928
REFERENCE/DOCKET NUMBER: UFLA:062\KIT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (713) 789-2679
INFORMATION FOR SEQ. ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 6280 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 988..1728
US-08-893-327-17

Query Match 12.1%; Score 981.8; DB 3; Length 6280;
Best Local Similarity 86.9%; Pred. No. 1.1e-210;
Matches 1162; Conservative 0; Mismatches 37; Indels 138; Gaps 2;

QY 809 ctgaagctcgtatcagcagcctgcagctgtgcctctcagttcagtcagcagcatcgtgtgttgc 868
Db 3044 CTGAGCTGCTCATACAGCTTCAGCTGCTGCTTCTAGTTCCAGCCATCTGTTGTGCT 3103
QY 869 cctcccccgtgcctctcttgcacccctggaagtgacacatcccaactgtccttccataaa 928
Db 3104 CTTCCCGCGTCTTACCTTACCTTGAAGAGTGCACACTCCACATGCTTCTTAATAAA 3163
QY 929 atagaagaattgcatgcattgtcgtatgtagtgcatttctatctcgggggtggtg 968
Db 3164 ATGAGGAATTTGCATGCTGCTGTGAGTGTCTATCTTCTGGGGGTGGGTGG 3223
QY 989 ggcaggaacgaaggggaagatctgggaagaataagcaagcagcagtcgtcgggga----- 1040
Db 3224 GGCAGACAGCAAGGGGAGGATTGGAGACAAATACAGGATGCTGGGAGAGATCTA 3283
QY 1041 ----- 1040
Db 3284 GGAACCCCTAGTATGAGATTGGCACTCCCTCTGCGCGCTCGCTACCTAGAGGC 3343
QY 1041 -----tgagtggtctctatggtcttctgaagcggaagaacag-- 1079
Db 3344 CGCCCGGCGCAAAACCGCGGCTGCGGCGACCTTGTGCGCGCCCTCAGTGAAGAGCG 3403
QY 1080 -----ctgataaag 1090
Db 3404 AGCGCCGACAGAGAGGAGTGGCCAAACCCCCCCCCCCCCCTGACAGCCCTGATTAATG 3463
QY 1091 aatcgccaacgcgcgggggaagcggttgcgtatagtcgagcttccgctcctcgcgt 1150
Db 3464 AATCGGCCAAACGCGGGGAGAGGGTTTGGCTATTGGCGCTCTTCGCTTCCCGCT 3523
QY 1151 cactgactcgtgcgtcgtggttgcgttgcgtgcggagcggtatcagctcaactcaagcg 1210
Db 3524 CACTGACTCGCTCGCTCGCTGCTGCTGCGCTGCGGCGATCAGCTACACTCAAGAGC 3583
QY 1211 ggtataaggtatcacagaatcagggagataacgaggaagaagaatgtgagcaaaag 1270

Db 3584 GGTATACGGTTATCCACAGATCAGGGGATACGCGAGAAAACATGTGACGAAAAAG 3643
QY 1271 ccagcaaaagccaggaacccgtaaaagccgctgtgctgttccatagagctcg 1330
Db 3644 CCACGAAAAAGGCCAGAACCTTAAAAAGCCGCTGCTGCTGCTTTTCCATAGGCTCCG 3703
QY 1331 ccccccctgaagagatcaaaaatcgacgctcaagtcagagtgagcgaaacccgacag 1390
Db 3704 CCCCCCTGACGAGCATCAAAAAATCGACGCTCAAGTCAGAGTGCGCAAAACCCGACAG 3763
QY 1391 actataagatacagagtggttccccctggagagctccctgtgctgtctctgttcgac 1450
Db 3764 ACTATTAAGATACCAAGGCGTTTCCCTCGAAGCTCCCTCGCGCTCTCTGTTCCAC 3823
QY 1451 cctgcgcttacccgatacctcgtccgcttcccttccttcggaaacgagcgcttccca 1510
Db 3824 CCTGCGGTTACCGGATACCTGCTCGCTTCTCCTTCGGGAACGCTGGGCTTTCACA 3883
QY 1511 atgtcagcgtgtatgtatctcaagttcggtgtagtgctgctcacaagctggagctgt 1570
Db 3884 ATGCTACGCTGTAGGTATCTCAGTTGCGGTAGTGTGCTGCTGCTGCTGCTGCTGCT 3943
QY 1571 gcaagaaaccccgcttaagcccgacgctgctgcttaccggttaactctgtcttgatc 1630
Db 3944 GCACGAACCCCGTTCAGCCCGACCGCTGCGCTTATCCGTAACCTTCTGAGTCC 4003
QY 1631 caaccgagtaagacagactatcgcacactgacagcaacacgtaacagattagcag 1690
Db 4004 CAACCCGGTAAAGACACAGCTTATGCGCACTGCGACACCACTGGTAACAGATTAGCAG 4063
QY 1691 agcagagatgtagcggtgtctacagagttcttgaagtgtgtgacctaaactacgctacac 1750
Db 4064 AGCAGGATATAGCGCGGTGCTACAGAGTTCTGAAGGTGGGCTTACTACGCTACAC 4123
QY 1751 tagaagagacgatttggtatctgctgctgctgctgaagcaattaccctcgaaaaaggt 1810
Db 4124 TAGAAGGACATATTTGGTATCTCGCTGCTGCTGCTGAAGCACTTACCTTCGAAAAAGAGT 4183
QY 1811 tggtagctctgtatccgcaacaacaacacacgctgtagcggtgtttttgtttgtaa 1870
Db 4184 TGGTAGCTCTTGATCCGGCAAAACCAACACGCTGTGTAGCGGTGTTTGTGTTGCAA 4243
QY 1871 gcagcagattacgcgcagaaaaaagatcacaagaagatccttgcattcttctacgg 1930
Db 4244 GCACGACATTCACGCGCAGAAAAAAGATCTCAAGAAATCTTGTGATCTTTCTACGCG 4303
QY 1931 gtctgaagctcagtggaagcaaaactacgcttaaggaattgtgtcagacgagatacat 1990
Db 4304 GTCTGACGCTAGTGGAACGAAACTCAGTTAAGGGATTTTGTGTCATGAGATTATCAAA 4363
QY 1991 attgaatgatttaga 2007
Db 4364 AAGGATCTTCACCTAGA 4380

RESULT 6
US-08-893-327-19
; Sequence 19, Application US/08893327
; Patent No. 6020192
; GENERAL INFORMATION:
; APPLICANT: Zoletukhin, Sergei
; APPLICANT: Hauswirth, William W.
; APPLICANT: Muzyczka, Nicholas
; TITLE OF INVENTION: Humanized Green Fluorescent Protein
; TITLE OF INVENTION: Genes and Methods
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P. O. Box 4433
; CITY: Houston
; STATE: TX
; COUNTRY: USA
; ZIP: 77210-4433

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/893,327
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/588,201
FILING DATE: 18-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Kitchell, Barbara S.
REGISTRATION NUMBER: 33,928
REFERENCE/DOCKET NUMBER: DPLA-062\KIT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (713) 789-2679
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 6280 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 988..1728
US-08-893-327-19

Query Match 12.1%; Score 981.8; DB 3; Length 6280;
Best Local Similarity 86.9%; Pred. No. 1.1e-210;
Matches 1162; Conservative 0; Mismatches 37; Indels 138; Gaps 2;

QY 809 cttagagctcgatgatacagctcgcagctgtgcttctagttgccagcactgtttgttgc 868
Db 3044 CTAGAGCTCGCTGATACACCTCGACTGCTTCTAGATTGCGACGACCATCTGTTTCC 3103
QY 869 cctccccgcttccttccttgcacccctggaaagtgccacccacactgtcttccaaataa 928
Db 3104 CTTCCCGGTGCTTCTTGAACCTGGAAGGTGCCACTGCTCCTTCTCTATATAAA 3163
QY 929 atgagaaatgtacatcgatgtgtcgtatgtagtgcattctatcttcctggggggtggg 988
Db 3164 ATGAGGAATGTGACATGCTGATGTGAGTGTGATTCATTCTGCGGGGTGGGATCTA 3223
QY 989 ggcagagacagcaaggggagagatgtggaaagaacaatagcagcagatgtgtggaa 1040
Db 3224 GGCAGGACAGCAAGGGGAGGATTGGGAACAATAGCAGCATCTCTGGGAGAGATCTA 3283
QY 1041 ----- 1040
Db 3284 GGAACTCTAGTATGATGAGTTGGCCACTCCCTCTCTGCGCGCTGCTGCTCATGAGGC 3343
QY 1041 -----tgcggtgggtctctatgtgtcttctgagggcggaagaaacag-- 1079
Db 3344 CGCCCGGCAAAAGCCGGGGGCTCGGGGACCTTTGTGCTGCGGCTCAGTGAAGAGCAAG 3403
QY 1080 -----ctgcattaatg 1090
Db 3404 AGCGGCGAGAGAGGAGTGGCCAAACCCCGCCCGCCCGCCCGCTGACGCTGCAATTATG 3463
QY 1091 aatcgccaaagcgcggggaaagcggtttgctgatatggcgcttcttcggttctcgct 1150
Db 3464 AATCGCCAAAGCGCGGGGGAAGGGGTTTGCATTATGGGCGCTTCTCTCTCGCT 3523
QY 1151 cactgaactcgtcgctcggttcgttcggtcgtcgagcgagcgtatcagctcaactcaaaagc 1210
Db 3524 CACTGACTCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3583
QY 1211 ggtataacgttatccacagaatcaggggataacgcaaggaagacatgtgaagcaaaag 1270
|||||

Db 3584 GGTAAATACGGTATTCACAGATCAGGGATTAACGAGAAAGACATGTGAGCAAAAG 3643
Qy 1271 ccagcaaaagccaggaacccgttaaaagccggttgctgctgcttccatagctccg 1330
Db 3644 CAGCAAAAGCCAGGAAACCGTAAAAAGCCGGCTGCTGCTTTTTCATAGGCTCCG 3703
Qy 1331 ccccccgtgagcatcacaaaatcgaagctaaagtcagaggttgagcaaacccagag 1390
Db 3704 CCCCCGTGAGAGATCACAAAATGACGCTTAAGTCAAGAGTGCGGAAACCCGACAGG 3763
Qy 1391 actataaagatacgaagcgttcccccgtgaagctccctcgtgcgtctcctgttccgac 1450
Db 3764 ACTATAAAGATACAGAGCGTTTCCCTGGAAGCTCCCTGCGCTGCTGTTCCGAC 3823
Qy 1451 cctgcgcgttacgagatacctgtcgcgttctcccttcgggaagcgttgagcttctca 1510
Db 3824 CCGCGCGCTTACCGGATACCTGTCCGCTTCCCTTCCGGAAGCGTGGCGCTTCTCA 3883
Qy 1511 atgctacgctgtagatcctagctcgtgtgagctgctcgtcgaagctgagctgtgt 1570
Db 3884 ATGCTACGCTGTAGATATCTCAATTCCGTTGAGTGTGCTGCTCCAGGCTGGGCTGTGT 3943
Qy 1571 gcaagcaaccccccgttcagcccgccagcgttgagccttaccgtaactatcgtctgagtc 1630
Db 3944 GCAGGAACCCCGCTTACAGCCCGACCGCTGCGCTTATCCGTAAGTATGCTTGTGATC 4003
Qy 1631 caaccggttaagaacacgaactatcgcgaactgagcagcaacatgtaacagatlagcag 1690
Db 4004 CAACCCGGTAAGACACGACTTATCGCCACTGCGACAGCACCTGTAACAGATTAAGCAG 4063
Qy 1691 agcgaagctgtgagcgtgtgctacagagttcttgaagttgttgcccaactagcgtacac 1750
Db 4064 AGCGAGGTATGTAGCGGGTCTACAGAGTCTTGAAAGTGTGCTGCTACAGGCTACAC 4123
Qy 1751 tagaaggaacagtaattgtatctcgtcgtcgtgaagccagttacccttcggaaaagagt 1810
Db 4124 TAGAAGGACAGTATTTGTATCTGCGCTGTGTAAGCCAGTTACCTCGGAAGAGAGT 4183
Qy 1811 tggtagctcttgatccgcaaaccaaacacgcgtgtgtagcgtgtgttttttcttgcaa 1870
Db 4184 TGGTACTCTTGTATCCGCGGAAACCAACCGCTGTAGCGGTGTTTGTGTTGCAA 4243
Qy 1871 gcaagcagattacgcgagaaaaaagatctcaagaagatcccttatcttcttcaggg 1930
Db 4244 GCAAGCATTTACCGCGAGAAAAAGAGATCTCAAGAGATCTTGTATCTTCTACGGG 4303
Qy 1931 gctcgaagctcagctggaacgaacacacacgcttaaggaatttgcatagcagatacat 1990
Db 4304 GCTTGACGCTCAAGTGAACGAACATCAGCTTAAGGATTTGTATGATGATTAACAA 4363
Qy 1991 attgaatgatttaga 2007
Db 4364 AAGGATCTTCACCTAGA 4380

RESULT 7

US-08-343-401A-3
; Sequence No. 5661132
; Patent No. 5661132
; GENERAL INFORMATION:
; APPLICANT: Swain, William F
; APPLICANT: Macklin, Michael D
; APPLICANT: Eriksson, Eilof
; APPLICANT: Andree, Christophe
; TITLE OF INVENTION: Improved Wound Healing
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: PO Box 2113
; CITY: Madison
; STATE: WI
; COUNTRY: USA
; ZIP: 53701-2113

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/343,401A
FILING DATE: 22-NOV-1994
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Seay, Nicholas J
REGISTRATION NUMBER: 27,386
REFERENCE/DOCKET NUMBER: 11-229-9103-9
TELECOMMUNICATION INFORMATION:
TELEPHONE: 608-251-5000
TELEFAX: 608-251-9166
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 4283 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
IMMEDIATE SOURCE:
CLONE: PMR01630
FEATURE:
NAME/KEY: exon
LOCATION: 713..721
FEATURE:
NAME/KEY: exon
LOCATION: 981..1253
FEATURE:
NAME/KEY: CDS
LOCATION: join(713..721, 981..1253)
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 713..1049
US-08-343-401A-3
Query Match 12.0%; Score 973; DB 1; Length 4283;
Best Local Similarity 83.2%; Pred. No. 8.6e-209;
Matches 1216; Conservative 0; Mismatches 30; Indels 215; Gaps 1;
Qy 762 ctgagacatcactatgagggccctattctatagtgtaacctaaatgtagagctgcgtg 821
Db 1277 CTCGAGCATCATCTAGAGGGCCCTATTATATAGTGTACCTTAATGCTAGAGCTGCTG 1336
Qy 822 atgaagctgactgtgctctcttagttgcaagcagcatcgtgtgttgcctcccccgtgcc 881
Db 1337 ATCAGCCTCAGCTGTGCTCTTAGTTCAGCCATCTGTTGTGCCCCCCTCCCGTGC 1396
Qy 882 ttccctgacctggaaggtgacacccactgcttcccttaataaaggaattgc 941
Db 1397 TTCTTGACCTCGAAGGTGCTCACTCCCACTGTCTTCTTAATAAATAGGAATTCG 1456
Qy 942 atcgcatctgctgtagtgcattctatctcgtggtgtggtgtggtggtggtggtggtggt 1001
Db 1457 ATCGCATCTGCTAGTAGAGTGTATCTTATTTGGGGGTTGGGGGTTGGGGGTCAGAGCAA 1516
Qy 1002 gggggaagattggaagaacaatagcaggaacgtctggtggtggtggtggtggtggtggt 1061
Db 1517 GGGGGAAGATTGGGAAGACAATAGCAGGATGCTGGGGATGCGGTGGGCTCTATGMAAC 1576
Qy 1062 tgagggcgaaaga----- 1074
Db 1577 AGCTGGGGCTCGAGATGCAAGCTTGAATTTCTATAGTGTACCTTAATAGCTTGCGCT 1636
Qy 1075 ----- 1074
Db 1637 AATCATGTCATAGCTGTTTCTGTGTGAATTTGTTATCCGCTCACAAATTCCACACACACA 1696
Qy 1075 ----- 1074

[illegible]

```

US-08-445-265A-1
: Sequence 1, Application US/08445265A
: Patent No. 5697901
:
: GENERAL INFORMATION:
: APPLICANT: Eriksson, Elof
: TITLE OF INVENTION: GENE DELIVERY BY MICRONEEDLE INJECTION
: NUMBER OF SEQUENCES: 4
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Quarles & Brady
: STREET: 1 South Pinckney Street
: CITY: Madison
: STATE: WI
:
: COUNTRY: US
: ZIP: 53703
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentln Release #1.0, Version #1.30
:
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/445,265A
: FILING DATE:
: CLASSIFICATION: 604
:
: ATTORNEY/AGENT INFORMATION:
: NAME: Seay, Nicholas J
: REGISTRATION NUMBER: 27386
: REFERENCE/DOCKET NUMBER: 110229.91080
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 608-251-9166
: TELEFAX: 608-251-5000
:
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 4283 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: circular
: MOLECULE TYPE: other nucleic acid
: DESCRIPTION: /desc = "Plasmid DNA"
:
: FEATURE:
: NAME/KEY: CDS
: LOCATION: join(713..721, 981..1250)
:
: US-08-445-265A-1

```

Query Match	12.0%	Score 973:	DB 1:	Length 4283:
Best Local Similarity	83.2%	Pred. No. 8.6e-209:		
Matches 1216:	Conservative 0:	Mismatches 309:	Indels 215:	Gaps 1:
QY 762	ctcgcgcgatgcatactcaaggaggccctattcctatagctgaaccctaattgctagaagctcgctg	821		
Db 1277	CTCGAGCATATGCACTTAGAGGGGCCCTATTCTTAAGTGTCACTAAATVGTATAGAGCTGGCGTG	1336		
QY 822	atcaaacccgcgaactgagcctcttagctgcacaagcactgttcttgaccctcccgcgtgac	881		
Db 1337	ATCAACCTCTGAGATGAGCTTCTTAGTTGGCAGGCANCTGTGTTTCCCTCCCCCTGCC	1396		
QY 882	ttcccttgaccctcggaaggtgcaccaccactgctcccttcctaataaaatgagaaattgac	941		
Db 1397	TTTCCTTGACCCCTGGAGAGGTGCCACTCCACATGTCTTTCCTAATAAATGAGAAATTGCG	1456		
QY 942	atcgcgactgtctcgaatgaagtgtcatctatctcttccttgaggggatgagggtgcaggaacagaa	1001		
Db 1457	ATCGCATGTGTCTGATATAGTATGTCTCATTTCTATTCTGGGGGGTGGGGGTGGGGGACGACAGCAA	1516		
QY 1002	gggggaggaattcggaagaacaatagcaggcactgctgggagtcggttgggctctatagcttc	1061		
Db 1517	GGGGGAGGATTTGGGAAGACAAATAGCAGGCATGCTGGGGATGGCGTGGGCTCATGTGAACG	1576		
QY 1062	tgaggcggaaaga	1074		
Db 1577	AGCTGGGGGTGCAGCATGCAAGCTTGAGTATTCATATGTCACCTAATAATAGCTTGGCGT	1636		
QY 1075		1074		

|||||
Db 3829 GATGAGCCTGAGCTGCTTCTAGCTGCAACCATCTGTGTGCTCCCTCCCGCTGC 3888
Qy 881 ctctcttacccttggaaggtgacactccacactgtctcttcccaataaatggaatltg 940
Db 3889 CTTCCTTACCCCTGGAAAGGTGGCACTCCACCTGCTTCTCAATAAATGAGAAATTTG 3948
Qy 941 catcgatctgtctgtgagtgatcatctatctctgaggggtgaggtgaggaacga 1000
Db 3949 CATGCAATGCTGTAGTAGTGTCTATCTATCTGGGGGTGGGGTGGGCGACGACAGA 4008
Qy 1001 agggggaagatgaggaacaaatagcagcagcagcagcagcagcagcagcagcagc 1060
Db 4009 AGGGGAGGATGGGAGAGACATAGCAGGCATGCTGGGATGCGGTGCGCTATGGAAC 4068
Qy 1061 ctgagagggagaaaga----- 1074
Db 4069 CAGCTGGGGCTGATCGATGATGATGATGATGATGATGATGATGATGATGATGATG 4128
Qy 1075 ----- 1074
Db 4129 TCATGCTATAGCTGTCTCTCTGTGAATTTGTATCCGCTCACAATTCACACACATA 4188
Qy 1075 ----- 1074
Db 4189 CGAGCGGAGACATAAAGTAAAGCTGGGGGTGCTTAATGATGAGTAACTACATTA 4248
Qy 1075 ----- 1074
Db 4249 ATTGGTTGCGCTCACTGCGCCCTTTCAGTGGGGAACCTGTCTGTCGCTGCACTTAA 4308
Qy 1089 tgaatcgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1148
Db 4309 TGAATCGGCGCAACGGCGGGAGAGCGGTGCGATTTGGGCGCTCTCCGCTTCTCG 4368
Qy 1149 ctactgactgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 1208
Db 4369 CTCACGTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4428
Qy 1209 gggggaataaggttcccaataacagggggaataagcaggaagaagaacatgagaga 1268
Db 4429 GGGGTAATACGGTTATCCACAGAAATCAGGGGATTAACGCGAAGAAATGTAAGAAAA 4488
Qy 1269 gggcagcaaaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1328
Db 4489 GGGCAGCAAAAGCCAGACACCGTAAAGGCGCTTCTGCGCTTTTCCATAGCTC 4548
Qy 1329 cgccccctgacgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1388
Db 4549 CGCCCCCTGACGAGCATCACAATAATGACGCTCAAGTCAGAGGTGCGGAAACCGACA 4608
Qy 1389 ggaactataagatlacagagcgttcccccctggaagcctccctgctgagcctctgttcg 1448
Db 4609 GGACTATAAAGATACAGCGCTTCCCTCGGAAGCTCCCTCGTGCGCTCTCTGTTCCG 4668
Qy 1449 accctgacgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1508
Db 4669 ACCCTGCGCTTACCGGATACCTGTCCGCTTCTCTTCCGGAAGCGTGCGCTTCT 4728
Qy 1509 caatgctcagcgttgaagatctcagctcagctcagctcagctcagctcagctcagctcag 1568
Db 4729 CAATGCTCAGCGTGTATGATCTCAAGTTGGGTAGGTGTTGCTGCTCAAGCGGGCTGT 4788
Qy 1569 gtgacagaaaccccgcttaagccgagcagcagcagcagcagcagcagcagcagcagc 1628
Db 4789 GTGACAGAAACCCCGCTTACGCCCGACCCCTGCGCTTATCCGCTAATCATGCTTTGAG 4848
Qy 1629 tccaacccgtaagacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1688
Db 4849 TCCAACCCCGTAAGACAGCACTTATGCGCAGTGGCGACGCACTGTGTAAACAGGATTAC 4908
Qy 1689 agagcaggtatgtagcgtgtctacagagctcttgaatgtgtgcttaactacgctac 1748
|||||

Db 4909 AGAGCAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGTGGCTTAACCTACGCTAC 4968
Qy 1749 actagaagacagatattgtaatctgctgctgctgctgctgctgctgctgctgctgctg 1808
Db 4969 ACTAAGAGCAGATATTGTGTATCTGCGCTCTGCTGAGCCAGTACTCTTGGAAGAA 5028
Qy 1809 gtgtgagctcttgaatccggaacaaacacacagcagcagcagcagcagcagcagcagc 1868
Db 5029 GTTGTAGCTTGTATCGGCAACCAACACCGCTGATGCGGTGCTGCTGCTGCTGCTG 5088
Qy 1869 aagcagcagatagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1928
Db 5089 AAGCAGCAGATTTACCGCAGAAAAAGAGATCTCAAGAGATCTTGAATCTTCTACG 5148
Qy 1929 gggctgacgctcagtggaacgaacacacagcagcagcagcagcagcagcagcagcagc 1988
Db 5149 GGGTCTGACGCTCAGTGTGAGCAACAAACCTCAGTAAAGGATTTGCTATGATTA 5208
Qy 1989 atattgaatgattaga 2007
Db 5209 AAAAGATCTCACCTAGA 5227
RESULT 12
US-08-667-769A-49
Sequence 49, Application US/08667769A
Patent No. 5783184
GENERAL INFORMATION:
APPLICANT: Ames, Robert S.
APPLICANT: Appelbaum, Edward R.
APPLICANT: Chaiken, Irwin M.
APPLICANT: Cook, Richard M.
APPLICANT: Gross, Mitchell S.
APPLICANT: Holmes, Stephen D.
APPLICANT: McMillan, Lynette J.
APPLICANT: Theisen, Timothy W.
TITLE OF INVENTION: Recombinant IL5 Antagonists Useful in
TITLE OF INVENTION: Treatment of IL5 Mediated Disorders
NUMBER OF SEQUENCES: 76
CORRESPONDENCE ADDRESS:
ADDRESSER: Smithline Beecham Corp./Corporate
STREET: P.O. Box 1539-UM2220
CITY: King of Prussia
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19406-0939
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/667,769A
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/17082
FILING DATE: 22-DEC-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/467420
FILING DATE: 06-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/470110
FILING DATE: 06-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/363131
FILING DATE: 23-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Sutton, Jeffrey A.
REGISTRATION NUMBER: 34,028
REFERENCE/DOCKET NUMBER: P50503
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-5024

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/637,647
FILING DATE: 19-JUN-1997
CLASSIFICATION: 536
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US/08/363,131
FILING DATE:
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Sulton, Jeffrey A.
REGISTRATION NUMBER: 34,028
REFERENCE/DOCKET NUMBER: P50282
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610 270-5024
TELEFAX: 610 270-5090
INFORMATION FOR SEQ ID NO: 49:
SEQUENCE CHARACTERISTICS:
LENGTH: 6285 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
US-08-637-647-49

Query Match 11.4%; Score 921.8; DB 3; Length 6285;
Best Local Similarity 82.9%; Pred. No. 2,9e-197;
Matches 1160; Conservative 0; Mismatches 27; Indels 212; Gaps 1;

QY 821 gatcagcctgactgctgcttctgtagtgcagcagcatctgttctgccccccgctgc 880
DB 3829 GATCAGCCTGAGCTGCTCTTCTAGTGCAGCATCTGTTGCCCCCCCCCGTGC 3888
QY 881 ctctctgagccctggaaaggggcccacatccatgctccttctctaataagaaattg 940
DB 3889 CTCTCTGAGCCCTGGAAAGGGGCCACTCCACTGCTCTTCTTAATAAATGAGAAATG 3948
QY 941 catcgactgtctgtagttagtgcattctctctgagggtgagggtgaggcagacagca 1000
DB 3949 CATCGACTGTGTGAGTAGTGTATCTATCTCTGAGGGGGGTGGGGCAGACAGCA 4008
QY 1001 agggggagagctgggaaagacatagcagcagctgaggatgctgctctctctgctt 1060
DB 4009 AGGGGAGAGCTTGGGAGACAAATAGCAGCATGCTGGGATGCGGTCTATGGAAC 4068
QY 1061 ctgaggggaaaga----- 1074
DB 4069 CAGCTGGGGCTCGATCGATGATGACTGCGGGCGCGATCCCGTCGAGAGCTTGGCGTAA 4128
QY 1075 ----- 1074
DB 4129 TCATGTCATAGCTGTTCTCTGTGTGAATTTGTTATCCGCTCACAATTCCACACAATA 4188
QY 1075 ----- 1074
DB 4189 CGAGCCGGAAGCATAAAGTGTAAAGCTGGGGTGCCTAATAGTAGCTAACTCATTTA 4248
QY 1075 -----acagcgtgcatlaa 1088
DB 4249 ATTGCGTGGCTCATCTGCCCGCTTTCACATCGGGAAACCTGTCGTGTCAGCTCATTTAA 4308
QY 1089 tgaatcgccaacagcgcgggggagagcggttgcgtatctgggagccttcgcctctcg 1148
DB 4309 TGAATCGCCAAAGCGCGGGGAGAGGGGTTTGGGTATTTGGGCGCTCTTCCGCTTCTCG 4368
QY 1149 ctcaatgactgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 1208
DB 4369 CTCACTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4428

QY 1209 gcggtatacgtttatccacagaatcagaggagatacgcaggaagaaacatgtgagcaaa 1268
DB 4429 GCGGTATACGGTATTATCCAGAAATCAGGGGATTAACGAGAGAAAGAACTGTGACAAA 4488
QY 1269 ggcagcaaaaagggcaggaacgcgtlaaaaagcgcgctgtgctgttctccatagctc 1328
DB 4489 GGCAGCAAAAAGGCGCAGAAACCGTAAAAAGGCCGCTGTGCTGCGGTCTTTCATAGCTC 4548
QY 1329 cggcccccctgacgagatcaaaaaatcgaagcctcaagtcgaagtggtgcgaaccgaca 1388
DB 4549 CGCCCCCTGACGAGCATCAAAAAATCGACGCTCAAGTCAAGAGTGGCGGAACCGACA 4608
QY 1389 ggaatataaagataccagagcgcttcccccctggaagctcccccctgagcctcctctccg 1448
DB 4609 GGAATATAAAGATACCAAGCGCTTCCCTCGAAGCTCCCTGCGCTCTCTCTGTTCCG 4668
QY 1449 accctgcgcttaccgagatacctgtccgcttctcccttccttcgggaaacgltgagccttct 1508
DB 4669 ACCCTGCGCGTTTACCGGATACCTGCTCCGCTTCTCCCTTCCGGAAGCGTGGCGCTTCT 4728
QY 1509 caatgctcaagctgtgagttatcagttcagttcagttcagttcagttcagttcagttc 1568
DB 4729 CAATGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCTGCTCCAACTGGGCTGT 4788
QY 1569 gtgagcaaacccccctgtaagcccgagcgcgtgagccttatccggttaactacgtctgag 1628
DB 4789 GTGCAAGAACCCCCCTGTACACCCGACCGCTGCGCTTATCCGTTACTATGCTTGTAG 4848
QY 1629 tccaaaccggtlaagacacagacttaccgaactggaagcagcagcagcagcagcagcagcagc 1688
DB 4849 TCCAAACCGGTAAACACAGACTTATCCGCACTGACAGCAGCAGCAGCAGCAGCAGCAGCAGC 4908
QY 1689 agagcaggtatgtaggcgtgtgctacagagttctgaagtggtgagccttaactacgctac 1748
DB 4909 AGAGCAGGATGTAGGCGCGGCTACAGAGTCTTGAAGTGTGCTGCTTACACTACGCTAC 4968
QY 1749 actaagaagacagatgtgtgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 1808
DB 4969 ACTAAGAGACAGATTTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5028
QY 1809 gtgtgactcttctgctcgc 1868
DB 5029 GTTGTAGCTCTTATATCCGGAACAAACACCGCTGTGACGGTGTGTTTTTGTGTTGC 5088
QY 1869 aagcagcagattacgcgcagaaaaaagatctcaagaagatcttgcattcttctacg 1928
DB 5089 AAGCAGCAGATTACGCCAGAAAAAAGATCTCAAGAAAGATCTTGTATCTTCTACG 5148
QY 1929 gggctcagcgtcagtggaacgaaactcaactcaaggaattgtgcatgagagagatc 1988
DB 5149 GGGCTCAGCCTCAGTGAACGAAACTCAGTTAAGGATTTGTGATGATGATATCA 5208
QY 1989 aatattgactataga 2007
DB 5209 AAAAGATCTTCACTAGA 5227

RESULT 15
PCT-US95-17082A-49
Sequence 49, Application PC/TUS9517082A
GENERAL INFORMATION:
APPLICANT: Ames, Robert S.
APPLICANT: Appelbaum, Edward R.
APPLICANT: Chaiken, Irwin M.
APPLICANT: Cook, Richard M.
APPLICANT: Gross, Mitchell S.
APPLICANT: Holmes, Stephen D.
APPLICANT: McMillan, Lynette J.
APPLICANT: Theisen, Timothy W.
TITLE OF INVENTION: Recombinant IL5 Antagonists Useful in
TREATMENT OF INFECTIONS: Treatment of IL5 Mediated Disorders
NUMBER OF SEQUENCES: 76
CORRESPONDENCE ADDRESS:

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 3, 2001, 15:21:54 ; Search time 5634.94 Seconds
(without alignments)
13613.261 Million cell updates/sec

Title: US-09-778-516-1

Perfect score: 8115

Sequence: 1 gatgtacggcgccagatatac.....gcgtgtcgtacgtacgagtc 8115

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues

Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

```
EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*
12: gb_est12:*
13: gb_est13:*
14: gb_est14:*
15: gb_est15:*
16: gb_est16:*
17: gb_est17:*
18: gb_est18:*
19: gb_est19:*
20: gb_est20:*
21: gb_est21:*
22: gb_est22:*
23: gb_est23:*
24: gb_est24:*
25: gb_est25:*
26: gb_est26:*
27: gb_est27:*
28: gb_est28:*
29: gb_est29:*
30: gb_est30:*
31: gb_est31:*
32: gb_est32:*
33: gb_est33:*
34: gb_est34:*
35: gb_est35:*
36: gb_est36:*
37: gb_est37:*
38: gb_est38:*
39: gb_est39:*
40: gb_est40:*
41: gb_est41:*
42: gb_est42:*
43: gb_est43:*
44: em_esthum10:*
45: em_esthum11:*
46: em_esthum12:*
47: em_esthum13:*
48: em_esthum14:*
49: em_esthum15:*
50: em_esthum16:*
51: em_esthum17:*
52: em_esthum18:*
53: em_esthum19:*
54: em_esthum20:*
55: em_esthum21:*
56: em_esthum22:*
57: em_esthum23:*
58: em_esthum24:*
59: em_esthum25:*
60: em_esthum26:*
61: em_esthum27:*
62: em_esthum28:*
63: em_estlin1:*
64: em_estlin2:*
65: em_estlin3:*
66: em_estlin4:*
67: em_estlin5:*
68: em_estlin6:*
69: em_estlin7:*
70: em_estlin8:*
71: em_estlin9:*
72: em_estlin10:*
73: em_estlin11:*
74: em_estlin12:*
75: em_estlin13:*
76: em_estlin14:*
77: em_estlin15:*
78: em_estlin16:*
79: em_estlin17:*
80: em_estlin18:*
81: em_estlin19:*
82: em_estlin20:*
83: em_estlin21:*
84: em_estlin22:*
85: em_estlin23:*
86: em_estlin24:*
87: em_estlin25:*
88: em_estlin26:*
89: em_estlin27:*
90: em_estlin28:*
91: em_estlin29:*
92: em_estlin30:*
93: em_estlin31:*
94: em_estlin32:*
95: em_estlin33:*
96: em_estlin34:*
97: em_estlin35:*
98: em_estlin36:*
99: em_estlin37:*
100: em_estlin38:*
101: em_estlin39:*
102: em_estlin40:*
103: em_estlin41:*
104: em_estlin42:*
105: em_estlin43:*
106: em_estlin44:*
107: em_estlin45:*
108: em_estlin46:*
109: em_estlin47:*
110: em_estlin48:*
111: em_estlin49:*
112: em_estlin50:*
113: em_estlin51:*
114: em_estlin52:*
115: em_estlin53:*
116: em_estlin54:*
```

117: gb_est48:*
 118: gb_est49:*
 119: gb_est50:*
 120: gb_est51:*
 121: gb_est52:*
 122: gb_est53:*
 123: gb_est54:*
 124: gb_est55:*
 125: gb_est56:*
 126: gb_est57:*
 127: gb_est58:*
 128: gb_est59:*
 129: gb_est60:*
 130: gb_est61:*
 131: gb_est62:*
 132: gb_est63:*
 133: gb_est64:*
 134: gb_est65:*
 135: gb_est66:*
 136: gb_est67:*
 137: gb_est68:*
 138: gb_est69:*
 139: gb_est70:*
 140: gb_est71:*
 141: gb_est72:*
 142: gb_est73:*
 143: gb_est74:*
 144: gb_est75:*
 145: gb_est76:*
 146: gb_est77:*
 147: gb_est78:*
 148: gb_est79:*
 149: gb_est80:*
 150: gb_est81:*
 151: gb_est82:*
 152: gb_est83:*
 153: gb_est84:*
 154: gb_est85:*
 155: gb_est86:*
 156: gb_est87:*
 157: gb_est88:*
 158: gb_est89:*
 159: gb_est90:*
 160: gb_est91:*
 161: gb_est92:*
 162: gb_est93:*
 163: gb_est94:*
 164: gb_est95:*
 165: gb_est96:*
 166: gb_est97:*
 167: gb_est98:*
 168: gb_est99:*
 169: gb_est100:*
 170: gb_est101:*
 171: gb_est102:*
 172: gb_est103:*
 173: gb_est104:*
 174: gb_est105:*
 175: gb_est106:*
 176: gb_est107:*
 177: gb_est108:*
 178: gb_est109:*

190: gb_est110:*
 191: gb_est111:*
 192: gb_est112:*
 193: gb_est113:*
 194: gb_est114:*
 195: gb_est115:*
 196: gb_est116:*
 197: gb_est117:*
 198: gb_est118:*
 199: gb_est119:*
 200: gb_est120:*
 201: gb_est121:*
 202: gb_est122:*
 203: gb_est123:*
 204: gb_est124:*
 205: gb_est125:*
 206: gb_est126:*
 207: gb_est127:*
 208: gb_est128:*
 209: gb_est129:*
 210: gb_est130:*
 211: gb_est131:*
 212: gb_est132:*
 213: gb_est133:*
 214: gb_est134:*
 215: gb_est135:*
 216: gb_est136:*
 217: gb_est137:*
 218: gb_est138:*
 219: gb_est139:*
 220: gb_est140:*
 221: gb_est141:*
 222: gb_est142:*
 223: gb_est143:*
 224: gb_est144:*
 225: gb_est145:*
 226: gb_est146:*
 227: gb_est147:*
 228: gb_est148:*
 229: gb_est149:*
 230: gb_est150:*
 231: gb_est151:*
 232: gb_est152:*
 233: gb_est153:*
 234: gb_est154:*
 235: gb_est155:*
 236: gb_est156:*
 237: gb_est157:*
 238: gb_est158:*
 239: gb_est159:*
 240: gb_est160:*
 241: gb_est161:*
 242: gb_est162:*
 243: gb_est163:*
 244: gb_est164:*
 245: gb_est165:*
 246: gb_est166:*
 247: gb_est167:*
 248: gb_est168:*
 249: gb_est169:*
 250: gb_est170:*
 251: gb_est171:*
 252: gb_est172:*
 253: gb_est173:*
 254: gb_est174:*
 255: gb_est175:*
 256: gb_est176:*
 257: gb_est177:*
 258: gb_est178:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Db 732 CACCCGGTAAGACAGCATTAATGCGCCACTGGACAGCCACTGTGTAACAGATTAGAG 791
Qy 1691 agcagatgtagttagcggatgtctacagattcttgaagtgatggtgagcctaactacggtacac 1750
Db 792 AGCGAGGATATAGGCGGTGCTACAGAGTTCTTGAATGGTGCGCTTAACCTACGCTACAC 851
Qy 1751 tagaagacagatttggatctgcgtctgctgaagcgaattacccitcggaanaaagt 1810
Db 852 TAGAAGACAGATATTGTTATCTGCGCTCTGCTGAAGCAGATTACTTCGGAAGAAAGT 911
Qy 1811 tggtagctcttgatcgcgcaacaacacacgcgtgtagcggt 1853
Db 912 TGGTAGCTTGTATCGGCAAAACACCGCTGTAGCGGT 954

RESULT 2

AU081137 1067 bp mRNA EST 15-NOV-1999
LOCUS AU081137 Oncorhynchus mykiss Kidney infected by infectious
DEFINITION hematopoietic necrosis virus Oncorhynchus mykiss cDNA clone K12,
mRNA sequence.
ACCESSION AU081137 GI:6431485
VERSION AU081137.1 GI:6431485
KEYWORDS EST.
SOURCE rainbow trout.
ORGANISM Oncorhynchus mykiss
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei;
Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
REFERENCE 1 (bases 1 to 1067)
AUTHORS Sakai,M. and Kono,T.
TITLE The EST analysis of kidney and gill cells rainbow trout,
Oncorhynchus mykiss, infected with infectious hematopoietic
necrosis virus
JOURNAL Unpublished (1999)
COMMENT Contact: Masahiro Sakai
Faculty of Agriculture
Miyazaki University
1-1 nishi gakuenkibanadai, Miyazaki, Miyazaki 889-2192, Japan
Email: aob208@cc.miyazaki-u.ac.jp.

FEATURES
Source Location/Qualifiers

1..1067
/organism="Oncorhynchus mykiss"
/db_xref="taxon:8022"
/clone="K12"
/clone_lib="Oncorhynchus mykiss Kidney infected by
infectious hematopoietic necrosis virus"
/tissue_type="Kidney infected by infectious hematopoietic
necrosis virus"

BASE COUNT 235 a 292 c 288 g 250 t 2 others
ORIGIN

Query Match 9.5%; Score 772.6; DB 107; Length 1067;
Best Local Similarity 98.6%; Pred. No. 1.2e-204;
Matches 800; Conservative 0; Mismatches 9; Indels 2; Gaps 2;

Qy 1076 ccagctgcatatgaatcgccaacgcgggaggaagcggttcgagttggggcgctc 1135
Db 259 CCAAGTGCATTATTAATCGCCAAACGCCGGGAGAGCGGTTTCGTAATGGCGCTC 318
Qy 1136 ttccgcttctcgtcactgaactgctgctgctgctgctgctgctgctgctgctgctgct 1195
Db 319 TTCGCTTCTCGTCACACGACGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTC 378
Qy 1196 agctactcaaaagcggtgaatcggtatccacaagaatcaggaggataacgcaggaaga 1255
Db 379 AGCTCACTCAAAAGCGGCTAATCGGTTATCCACAGATCAGGGGATTAACGCAAGAAAGA 438
Qy 1256 catgtagagcaaaagccagcaaaagccaggaacggttaaaggccgctgctgctgct 1315
Db 439 CATGTGAGCAAAAGCGCACCAAAAGCGCAGAAACGTTAAAGAGCCGCGTGTGCGCTIV 498

Qy 1316 ttccataggtccgcccccctgacagcatcaaaaaatcgaagctcaagtcaagtg 1375
Db 499 VTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAAATGCACCTCAAGTCAAGAGTG 558
Qy 1376 gggaaaccggacagactatagaatatacgaagcgtttcccccctggaagctccctgctg 1435
Db 559 GGAAGACCCGAGAGAGCTATTAAGATACAGGCGTTTCCCTGGAGAGTCCCTGCTGCG 618
Qy 1436 ctctccttccgaccctgcgcgcttaccgatacctgtgcgcttctcccttcgggaag 1495
Db 619 CTCTCCTGTGTCGACCTGCGCCTTACCGGATACCTGTGTCGCTTTCCTTCGGGAAG 678
Qy 1496 cgtgagcttctcgaatgctacgctgtagtatactcaagttcgtgtgtgctgctc 1555
Db 679 CGTGCGGCTTCTCATTAGCTACGCTGTAGTATCTCAGATTCTGATGTTGCTGCTC 738
Qy 1556 caagctgagctgtgtgcagcaaacccccctgtaagccagccgtcgccttaacggtaa 1615
Db 739 CAGCTGGGCTGTGTGCAGAACCCCGCTTACGCCCGACCGCTGCTTATCCGGTAA 798
Qy 1616 ctatgcttgatgcacaacccgtaagaacagactatgcacactgagcagcagcactg 1675
Db 799 CTATGCTTGTAGTCCAAACCCGTTAAGACAGATTATGCGCAGTGGCAGCCACTGG 858
Qy 1676 taacagattagcagaagcaggtatgtaggggtgctacagaagttcctgaagtgtg 1735
Db 859 TAACAGGATTACAGACGCGATATGTAGCGGCTCTACAGAGTTCTTGAAGTGTGGCC 918
Qy 1736 taactaagctactactagaagaacagattgtgatactgcgtcgtcgtcaagcgttac 1795
Db 919 TAACTAAGGCTTACCTAGAGACAGATATTGTGATCTGCGCTGTGCTAACCAGATTAC 978
Qy 1796 ctgcgaaaaagatgtgtgtagctcttgatccgycgcaaaaacacccgcgtgtgtagcg 1855
Db 979 CTTGGAAGAAAGATGTGGTAGCTTTGATCCGCAAAACA-CAACGCTGTAGGCTTT 1037
Qy 1856 ttttttgcgaagcagcaattacgctc 1886
Db 1038 TTTTGTGTTTGC-AGCAGCAGATTACGGCG 1067

RESULT 3

AU081124 1089 bp mRNA EST 15-NOV-1999
LOCUS AU081124 Oncorhynchus mykiss Kidney infected by infections
DEFINITION hematopoietic necrosis virus Oncorhynchus mykiss cDNA clone KG'12,
mRNA sequence.
ACCESSION AU081124 GI:6431472
VERSION AU081124.1 GI:6431472
KEYWORDS EST.
SOURCE rainbow trout.
ORGANISM Oncorhynchus mykiss
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei;
Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.

REFERENCE 1 (bases 1 to 1089)
AUTHORS Sakai,M. and Kono,T.
TITLE The EST analysis of kidney and gill cells rainbow trout,
Oncorhynchus mykiss, infected with infectious hematopoietic
necrosis virus

JOURNAL

Unpublished (1999)
COMMENT Contact: Masahiro Sakai
Faculty of Agriculture
Miyazaki University
1-1 nishi gakuenkibanadai, Miyazaki, Miyazaki 889-2192, Japan
Email: aob208@cc.miyazaki-u.ac.jp.

FEATURES
Source Location/Qualifiers

1..1089
/organism="Oncorhynchus mykiss"
/db_xref="taxon:8022"
/clone="KG'12"
/clone_lib="Oncorhynchus mykiss Kidney infected by
infectious hematopoietic necrosis virus"

TITLE	Anopheles gambiae pilot gene discovery project: identification of mosquito innate immunity genes from expressed sequence tags generated from immune-competent cell lines					
JOURNAL MEDLINE COMMENT	Proc. Natl. Acad. Sci. U.S.A. 97 (12), 6619-6624 (2000) 20300950 Contract: Dimopoulos G Fotis C. Kafatos Laboratory European Molecular Biology Laboratory Meyerhofstrasse 1, 69117 Heidelberg, Germany. Location/Qualifiers 1..700 /organism="Anopheles gambiae" /strain="4A r/f" /db_xref="taxon:7165" /clone="4A3A-P8A12" /clone.lib="Anopheles gambiae immune competent 4A3A" /cell.line="immune competent 4A3A" /lab.host="E. coli DH10B" /note="Vector: pUT3D-Pac (Pharmacia) with a modified polylinker; Site_1: EcoRI; Site_2: NotI; sequenced forward priming site which reads from the 3' end of the cDNA. The 4A3A is a directionally cloned and normalized cDNA library that was constructed from the 4A3A cell line oligo-T primed cDNA according to: Bonaldo, Lennon & Soares (1996) : Normalization and Subtraction: Two Approaches To Facilitate Gene Discovery, Genome Research 6, 791-806."					
BASE COUNT	167 a	196 c	189 g	146 t	2 others	
ORIGIN						
Query Match	8.3%	Score 677.4:	DB 104:	Length 700:		
Best Local Similarity	98.8%;	Pred. No. 4.8e-178;	Matches 681;	Conservative 0;	Mismatches 8;	Indels 0;
Gaps	0;					
OY 1163	ggcgctcgcgtcgttgcgctgcygcgcgcgcgatccagctcaactcaagaagcggtatcagtt	1222				
Db 12	GGTTGCGGTCTTCCGCTGCAGCGGATACACTCCTCAAAAGCGGTATACGGTT	71				
OY 1223	atccacagatcacagggaataacgcagaaagaacatgttagcaaaaagccagcaaaagc	1282				
Db 72	ATCCACAGAATACGGGGAATACGCAGAAAGAACATGTGACCAAAAAGCCAGCAAAAGCC	131				
OY 1283	caggaaaccgtaaaaagccgcgtctgcgttgcttttccaiaagctccgcccttcagca	1342				
Db 132	CAGGAACCCTAAAAAGCGCGCTTGCTGCAGNNTTTCATAGCTCCGCCCCCTGACGA	191				
OY 1343	gatatcaaaaatcgaagctcaatgcatagtcagagtggtggagaaacccagacaagtataagata	1402				
Db 192	GCATACCAAAAATCACACGCTCAAAGTCAGAGGCGGGGAAACCCGACAGCACTTAAAGATA	251				
OY 1403	ccaggcgattccccctggagaagctccactcgttgtgcctctcctgtfctcgaaccttcgcgttac	1462				
Db 252	CCAGGCGTTTTCCCCCTGGAAAGCTCCCTCGTGGGCTCTCTGTTCGACCCCTGCTTAC	311				
OY 1463	cggataacctgcgccttctcccttcgcggaaagcgtyggcgctttctcaatgcatagcactgy	1522				
Db 312	CGGAAATACGTGCGGCTTCTCCTTCCTGGGAAGCGTGGGCGTTTCTCATATAGCTACGGCG	371				
OY 1523	tagtatcatcagttcgggtgtagtgctgcgtccgaagctgggctgtgtgcacgaacccc	1582				
Db 372	TAGGATATCAATTGGGTGTAGGTGTTGGCTCCAAAGCGGGCTGTGTGCAGAAACCCC	431				
OY 1583	ggttaagcccgaccgctgcgccttaccggtatcacatgctcttggttccaacccggttag	1642				
Db 432	CGTTACGCCGACCGCTGGCCTTAATCCGGTACATATGCTCTTGATGCCAACCCCGTAGG	491				
OY 1643	aacagactatgcagcactgycagcagccacactggtatacagatltaacagacgagtgatgt	1702				
Db 492	AACAAGCTTATGCGCACGTGCGACGACCACTGTAAACAGATTACAGACGACGAGTATGT	551				
OY 1703	aggcggtgtctacagagttcttgaagttgvtgacctactaagctactaagaagacagt	1762				
Db 552	AGCGGTGTATACAGAGTTCTTGAAGTGTGGGCTTAATACAGGCTACACTAGAAAGACGT	611				

QY	1763	atgttgatctctcgtctcgtcgaacccagttacctctcgaaagaagcttgtagctctg	1822
Db	612	ATTGTGATCTCGCTCTGCTGTAACCCAGTACCTTTCGGAAAAAGATTGTGACTCTTG	671
QY	1823	atccggcaacaacaccgcgtctgtagcg	1851
Db	672	ATCCGGCAACAACACCACCGCTGTAAGCG	700
RESULT	7		
LOCUS	AJ281552		
DEFINITION	4A3A-p6F11-F Anopheles gambiae immune competent 4A3A Anopheles	EST	30-JUN-2000
ACCESSION	AJ281552		
VERSION	AJ281552.1	GI:6929432	
KEYWORDS	EST.		
SOURCE	African malaria mosquito.		
ORGANISM	Anopheles gambiae		
REFERENCE	1 (bases 1 to 1070)		
AUTHORS	Dimopoulos, G., Casavant, T. L., Chang, S., Scheetz, T., Roberts, C. J., Donohue, M., Schultz, J. J., Benes, V., Bork, P., Ansoirge, W., Soares, M. B. and Kafatos, F. C.		
TITLE	Anopheles gambiae pilot gene discovery project: identification of mosquito innate immunity genes from expressed sequence tags generated from immune-competent cell lines		
JOURNAL	Proc. Natl. Acad. Sci. U.S.A.	97 (12),	6619-6624 (2000)
MEIDLIN	20300950		
COMMENT	Contact: Dimopoulos G Fotis C. Kafatos Laboratory European Molecular Biology Laboratory Meyerhofstrasse 1, 69117 Heidelberg, Germany. Location/Qualifiers 1. 1070		
FEATURES	source		
	1. 1070		
	/organism="Anopheles gambiae"		
	/strain="4A r/F"		
	/db_xref="taxon:7165"		
	/clone="4A3A-p6F11"		
	/clone.lib="Anopheles gambiae immune competent 4A3A"		
	/cell_line="Immune competent 4A3A"		
	/lab_host="E. coli DH10B"		
	/note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: EcoRI; Site_2: NotI; sequenced from forward priming site which reads from the 3' end of the cDNA. The 4A3A is a directionally cloned and normalized cDNA library that was constructed from the 4A3A cell line oligo-T primed cDNA according to: Bonaldo, Lennon & Soares (1996) : Normalization and Subtraction: Two approaches to Facilitate Gene Discovery, Genome Research 6, 791-806."		
BASE COUNT	263 a	283 c	255 g
ORIGIN	269 t		
Query Match	8.3%,	Score 672.6;	DB 104; Length 1070;
Best Local Similarity	97.3%;	Pred. No. 1.2e-176;	
Matches 684;	Conservative 0;	Mismatches 19;	Indels 0; Gaps 0;
QY	1305	ttgcgcgcgtttttccataagctccgccccctcgacgcgcatacacaataatcgagctca	1364
Db	1	TTTGCTGGCGGTTTTCATATAGGCTCCGCCCTCGACGACATCACAAAATCGAGCTCA	60
QY	1365	agtcagaggtgagcaaacccgcacagactataaagatcacagcgcgtttcccccctggaagc	1424
Db	61	AGTCAGAGGTGGCGAACC CGACAGACTATTAAGATACAGAGGGGTTTCCCCCTGGAAGC	120
QY	1425	tcctcgtgtagctctcctgttcgcgaacctgcgcgttaccgcgatatcctgtccgcctttctc	1484
Db	121	TCCCTTCGTCGCTCTCCCTGTTCCGCACCTCCGCGCTTACCGGATACCTGTCCGCTTTCTC	180

QY 1485 cctcggaaacgctgacgtcttctcaatgctcaacgctgtagtattcagttcagttcgtgac 1544
|||||
Db 181 CCTTCGGGAAGCGTGGCGCTTCTCATAGCTCAGCCTGAGTATCTCAGTTGGGTGAG 240
QY 1545 gtcgttcgtccaaagctggtgtgtgcaagaaccccccggttcagcccgagcgtgagcc 1604
|||||
Db 241 GTCGTCCTCCCAACGCTGGCGTGTGTGACGAAACCCCGCTTCACCCGACCGCTGCGCC 300
QY 1605 ttatccgtaactatcgtcttgagctcaaccccggtgaagaacagactttcggccattgca 1664
|||||
Db 301 TTATCCGTAACCTATCGTCTTGAGTCCAAACCCGGTAAGACAGACTTATCCGCTAGGCA 360
QY 1665 gcagccactggttaacaagattagcagagcgaagtatgttagcggtgctacagagttctg 1724
|||||
Db 361 GCAGCCACTGGTAACAGATTAGCAGAGCAGAGTATGTAGCGGTGCTACAGATTCTTG 420
QY 1725 aagtggtgcttaactaagcgtacactagaagaagcagattatgtgtatctgctgctgtg 1784
|||||
Db 421 AAGTGTGGCCTTAACCTACGCGCTACACTAGAAACAGATTTGTATCTGCCCTCTGCTG 480
QY 1785 aagccagttactcgcggaagaagattgtagctcttgatcccggaacaacacacacgct 1844
|||||
Db 481 AAGCCAGTACTCTTGGAAGAAAGATTGGTACTCTTGATCCGGAACAAACACCGCT 540
QY 1845 gtagcagtggtttttttgttctgcaagcagcagattacgcgcgagaaaaaagatctcaa 1904
|||||
Db 541 GGTAGCGGTGGTGTGTGTGTTCGCAAGCAGCAGATTACGCGCAGAAAAAAGATCTCA 600
QY 1905 gaagatccttgatcttcttctcaggggtcctgacgtcagtggaacgaacacacgttaa 1964
|||||
Db 601 GAGATCTTGTGATCTTCTTCTACGGGGCTGTGACGCTCAGTGAAGCAAAACCTCACGTAA 660
QY 1965 gggatttgatcagcagcagacacatttgatattataga 2007
|||||
Db 661 GGGATTTGGTCATGATGATATCAAAAAAGATCTTCACCTAGA 703

RESULT 8
AJ281437
LOCUS 703 bp mRNA EST 30-JUN-2000
DEFINITION 4A3A-P4C3-F Anopheles gambiae immune competent 4A3A Anopheles
gambiae cDNA clone 4A3A-P4C3, mRNA sequence.
ACCESSION AJ281437
VERSION AJ281437.1 GI:5929317
KEYWORDS EST.
SOURCE African malaria mosquito.
ORGANISM Anopheles gambiae
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae
; Anopheles.
REFERENCE 1 (bases 1 to 703)
AUTHORS Dimopoulos, G., Casavant, T. L., Chang, S., Schaeitz, T., Roberts, C.,
Donohue, M., Schultz, J., Benes, V., Botk, P., Ansoerge, W., Soares, M. B.
and Kafatos, F. C.
TITLE Anopheles gambiae pilot gene discovery project: identification of
mosquito innate immunity genes from expressed sequence tags
generated from immune-competent cell lines
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (12), 6619-6624 (2000)
MEDLINE 20300950
COMMENT Contact: Dimopoulos G
Fotis C. Kafatos Laboratory
European Molecular Biology Laboratory
Meyerhofstrasse 1, 69117 Heidelberg, Germany.
FEATURES
Source location/qualifiers
1..703
/organism="Anopheles gambiae"
/strain="4A r/r"
/db_xref="taxon:7165"
/clone="4A3A-P4C3"
/clone_lib="Anopheles gambiae immune competent 4A3A"
/cell_line="Immune competent 4A3A"
/lab_host="E. coli DH10B"
/note="Vector: pT73D-Pac (Pharmacia) with a modified

polylinker; Site_1: EcoRI; Site_2: NotI; sequenced from
forward priming site which reads from the 3' end of the
cDNA. The 4A3A is a directionally cloned and normalized
cDNA library that was constructed from the 4A3A cell line
Oligo-T primed cDNA according to: Bonaldo, Lennon & Soares
(1996) : Normalization and Subtraction: Two Approaches To
Facilitate Gene Discovery, Genome Research 6, 791-806."

BASE COUNT 159 a 204 c 193 g 147 t
ORIGIN

Query Match 8.2%; Score 663; DB 104; Length 703;
Best Local Similarity 97.8%; Pred. No. 5.2e-174;
Matches 672; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1075 accaagtcattatgaatcgcgcaacgcgagcgagggagggcggtttggtattggcgct 1134
|||||
Db 17 ACTAAAGGGAATTAACCTTGCGGGCAACCGCGGGAAGAGCGGTTGCTATTGGACGCT 76
QY 1135 ctccgcttcctcgtctcaactgactcgtcgtcgctcgctcgctcgctcgagcggtat 1194
|||||
Db 77 CTTCCGCTTCCCTCGCTCACTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 136
QY 1195 cagctcactcaaaagcgytaacggttatccacgaatcaggggataacgcagaaaga 1254
|||||
Db 137 CAGCTCACCTCAAAAGCGGTAATATACGTTATCCACAGAAATCAGCGGATTAACGAGAAAGA 196
QY 1255 acatgtgcaaaaagccagcaaaaagccaggaacccgttaaaaagccggtgtcgagct 1314
|||||
Db 197 ACATGTGAGCAAAAGCCAGCAAAAGCCAGGACCGTAAAGGCCGCGTGTGCGCT 256
QY 1315 ttctcctaagcgtccgccccctgacgagcatcacaaaatcgacgtcctaagtcagagtc 1374
|||||
Db 257 TTTTCATAGGCTCCGCCGCCCTTACGAGCATCACAAAATGAGCGCTCAATCAGAGGT 316
QY 1375 ggcgaaacccgcagagactataaagatacagaggtttcccccgtggaagctccctcgtgc 1434
|||||
Db 317 GCGCAAAACCCGACAGAGACTATTAAGATACACAGCGCTTCCCTCGAAGCTCCCTGTC 376
QY 1435 gcttcctcgttcgcgacccctcgccgttaccgagtaactcgtcccttcctcccttgaga 1494
|||||
Db 377 GCTGCTCCTGTTCCGACCCCTGCGCTTACCGGATACCTGTGCCCTTTCTCCCTTGGGGA 436
QY 1495 gcttgagcgtcttctcaatgctcagcgtgtagttatctcagttcgtgtgtagtgcgtct 1554
|||||
Db 437 GCGTGGCGCTTCTCATAGCTCAGCCTGAGTATCTCAGTTCCGGGTAGTGCTGCTTCCGT 496
QY 1555 ccaagctgggctgtgtgcagaaccccccggttcagccgagccgctgcgcttaccgtta 1614
|||||
Db 497 CCNAGCTGGGCTGTGTGACGAAACCCCGCTTCACGCCGACCGCTGCCCTTATCCGGTA 556
QY 1615 actatgctcttgatgacacaccccggtgaagacagacttttcggccattgagagacacgt 1674
|||||
Db 557 ACTATGCTCTTATGTCACACCCGTTAAGACACGACTTATCGCACCTGGAGCAGCCACTG 616
QY 1675 gtaacagatagcagagcagagatgttagcgtgtgctacagagttcttgaagtgtgtgc 1734
|||||
Db 617 GTAACAGATTAGCAGAGGAGAGTATGTAGGGGCTGCTACAGATTCTTGAAGTGTGTGC 676
QY 1735 ctaactacggtcactacactagaagacag 1761
|||||
Db 677 CTAACTAGGGCTACACTAGAAAGACAG 703

RESULT 9
BF863156 757 bp mRNA EST 19-JAN-2001
LOCUS BF863156
DEFINITION 963041F03.x1 C. reinhardtii CC-1690, Stress condition I, normalized
lambda zap II Chlamydomonas reinhardtii cDNA, mRNA sequence.
ACCESSION BF863156
VERSION BF863156.1 GI:12253300
KEYWORDS EST.
SOURCE Chlamydomonas reinhardtii.

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
ORGANISM	Chlamydomonas reinhardtii			
REFERENCE	Eukaryota,Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales; Chlamydomonadaceae; Chlamydomonas.			
AUTHORS	Grossman,A., Davies,J., Federspiel,N., Harris,E., Hauser,C., Lefebvre,P., McDermott,J.P., Shriver,J., Sillflow,C. and Sten,D.			
TITLE	Analyses of the Chlamydomonas reinhardtii Genome: A Model, Unicellular System for Analyzing Gene Function and Regulation in Vascular Plants; project phase 3			
JOURNAL	Unpublished (2000)			
COMMENT	Contact: Charles Hauser DCMB Box 91000 Duke University Durham, NC 27708-1000 Tel: 919 613 8159 Fax: 919 613 8177 Email: chauser@duke.edu.			
FEATURES				
SOURCE	Location/Qualifiers 1..757 /organism="Chlamydomonas reinhardtii" /strain="CC-1690 wild type mt+ 219r" /db_xref="taxon:3055" /clone_lib="C. reinhardtii CC-1690, Stress condition I, normalized, Lambda Zap II" /note="Vector: pBluescript II SK-; Site.1: EcoRI; Site.2: XhoI; This library, constructed by John Davies and Jeffrey McDermott, combines cDNAs from CC-1690 cells grown to mid-log phase in TAP-N (30 min, 1hr, 4hr), TAP-S (30 min, 1hr, 4hr), TAP-P (4hr, 12hr, 24hr), NO3 to NH4 (30min, 1hr, 4hr) and NH4 to NO3 (30min, 1hr, 4hr). PolyA mRNA was purified from each sample, pooled and cDNA synthesized. The cDNA was directionally cloned into lambda zap II (Stratagene) in the EcoRI (5') and XhoI (3') sites. pBluescript II SK- plasmids were excised from the lambda Zap clones by superinfection with Exsist (Stratagene) phage. The library was normalized using method 4 described in Bonaldo et al (1996) Genome Research 6: 791-806."			
BASE COUNT	167 a 219 c 208 g 161 t 2 others			
ORIGIN				
Query Match	8.0%; Score 653.2; DB 170; Length 757;			
Best Local Similarity	95.3%; Pred. NO. 3e-171;			
Matches	684; Conservative 0; Mismatches 33; Indels 1; Gaps 1;			
QY	1035 tgggagatcgcggttgagctctatgctctctgagcggaagaacacagctgcatatgaatc 1094			
DB	34 TGGGTACCGGGGCCCGCGTTTCCGTGCGGGAACCTGTGTGCGACATCATTAATGATC 93			
QY	1095 ggcacaagcgcggggagagggcggttggtgatttggcgcccttcgcgttcctgcgtact 1154			
DB	94 GGCCAAACGGCGGGGAGGCGGCGTTTGGCATTTGGCCCTCTCCGTTCCCTCGCTCACT 153			
QY	1155 gactcgtcgcgctcgtctcgtctcgtctgcgtgcgcgagcgatcacagctcaacaaagcgcta 1214			
DB	154 GACTCGCTGCGCGTGGTGGTTCGCGTGGGCGGAGGGGATACGTCACATCAAAAGCGGTA 213			
QY	1215 atacggtatccacaagaatcaggggataacgcaggaagaacatgtgacaaagccag 1274			
DB	214 ATACGTTATTCACACAGATTCAGGGGATTAACGACGGAAGAACATGTGACAAAGGCCAG 273			
QY	1275 caaagagcagaagaaccgtaaaaaagccgcgttgcgcgggttttccataagctccgccc 1334			
DB	274 CAAAGAGCCAGGAACCGTAANAAGCCCGCGTGGCTGGGGTTTTCATTAAGGCTCCGCC 333			
QY	1335 cctgacgagcatcaaaaaatcgacgctcaagtcagaggttgcgagaaacgcagacagacta 1394			
DB	334 CCTGACGAGCATCACAAAATGACGCTCAAGTACAGAGGTGGCGAAGCCGACAGAGACTA 393			
QY	1395 taagaataccagggcttccccccttgagaaagctccctcgtgtgcctctcgtttccagacctg 1454			
DB	394 TAAAGATTCACAGGCGCTTTCCCTCGGAAGCTCCCTCGCGCTGTCTCCGTTCCGACCCGTG 453			

QY	1455	cgcgctaccggatacctgtcgcgccttctcccttcctgggaagcttgccgttctctaatgc	1514
Db	454	CCGCTTACCGGATACCTGTGTCGCCCTTTCCTCCCTCGGGAAGCGTGGCTTTCATAC	513
QY	1515	tcacgtctgaagatattcaagttcgcgtgtgtagtgcgttcgcgtcccaagcttgagctgtgcaac	1574
Db	514	TCACGCTGTAGGATCTCAGTTCCGGTGTAGTGCGTTCGCTCCACAGCTGGGCTGTGTGCAC	573
QY	1575	gaaccccccgcttcgaacccgcacgcgtgagccttatcccgtaactatgcgtcttgagtcacaac	1634
Db	574	GAACCCCCCGTTACAGCCCAACCGCTCCGCTTATCCGGTAACTATGCTTGAAGTCAAC	633
QY	1635	ccggttaagcacaagactattatcgccacatgagcagcagcagcagcgtgttaacaggtttgacagag	1694
Db	634	CCGGTAAAGACAGACTTATTCGCCACTGCGACGACGCCATCGTAAACGGATTAGCAACAGCG	693
QY	1695	aggtatgtatggaggtgctctacagatctctgaagttgtgcttcctaactacagctctacacta	1752
Db	694	AGGTATGTATGGCGGTCTACAGAGTTCTTGGAGTGTAGTACGTACACACGGCTACACTA	750
RESULT	10		
AL039459			
LOCUS	AL039459	734 bp	mRNA
DEFINITION	DKFZP43401210.s1.434 (synonym: htes3)	Homo sapiens	cDNA clone
ACCESSION	AL039459		
VERSION	AL039459.2	GI:5935158	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
AUTHORS	1 (bases 1 to 734)		
	Duesterhoeft,A., lauber,J., Mewes,H.W., Gassenhuber,J. and Wiemann		
	,S.		
	EST (Duesterhoeft, et al.)		
	Unpublished (1999)		
	On Jul 7, 1999 this sequence version replaced gi:5408507.		
COMMENT	Contact: Duesterhoeft A		
	MIPS		
	Am Klopfersplitz 18a D-82152 Martinsried, Germany		
	This is the 3' sequence of the clone insert		
	Clone from S. Wiemann, Molecular Genome Analysis, German Cancer		
	Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de;		
	sequenced by Qiagen (Hilden/Germany) within the cDNA sequencing		
	consortium of the German Genome Project.		
	No r1 sequence available.		
	This clone (DKFZP43401210) is available at the RZPD in Berlin.		
	Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059		
	Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.		
FEATURES	Location/Qualifiers		
source	1..734		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="DKFZP43401210"		
	/clone_lib="434 (synonym: htes3)"		
	/tissue_type="testis"		
	/dev_stage="adult"		
	/lab_host="DH10B"		
	/note="Vector: pSport1; Site_1: NotI; Site_2: SalI"		
BASE COUNT	172 a 199 c 194 g 161 t 8 others		
ORIGIN			
Query Match	7.7%; Score 621.4; DB 105; Length 734;		
Best Local Similarity	94.4%; Pred. No. 2.4e-162;		
Matches	685; Conservative 0; Mismatches 33; Indels 8; Gaps 4		
QY	1159	gcgtcgcctcggctgttcgcgtgcgagcgagcggtatcagctcaactcaagcgcgtaataac	1218
Db	9	CGCTCGCGTGGCTGCTTGGCTGCGCGCGCGATACCTCACTCACTCAAGCGGTAAATAC	68

OY	1219	ggatctccacgaatctcaagggaataacgaaggaagaacatgtgaagcaaaaagcccaagaa	1278
Db	69	GGTATTCACAGAAATCAGGGGATTAACGAGGAAAGAAACATGTAGCAAAAGGCCGCAAA	128
OY	1279	agggccaggaacccgtaaaaaggccgcgtctgcgtgcgtttccataagactccgcgccctg	1338
Db	129	AGGCCAGGAACCGTAAAAAGGCCGCGTTGCGTGGGCTTTTTCATAGGCTCGGCCCTTG	188
OY	1339	acgagcatcacaanaatcgcgtctcaagtcaggaaggtgagcaacccgaagactataa	1398
Db	189	ACGACATCACAAAAATCGACGCCTCAAGTCAGAGGTGGCGAAACCCACAGGACTATAAA	248
OY	1399	gataccaagcgcttcccccttgaaagctcccccgtgcgcgtccctccgttccgaacctgcgc	1458
Db	249	GATACACAGGCGTTTCCCTTGAAAGCTCCCTCGTGGCGCTCTCTGTTCGACCCCTGGCG	308
OY	1459	ttaccggaataccttgcgccttccctctcccttcgggaagcgtgcgttctcctaagtctac	1518
Db	309	TTACCGGATACCTGTCGCGCTTCTTCCTTCGGGAACGTCGGCGCTTTCATAGCTCAC	368
OY	1519	gctgtaggatctcaagtcagtgtagtgcgttcgctccaagctggtcgtgtgcaegaac	1578
Db	369	GCTGTAGGTATCTCAGTTGCGGTGTAAGTCGTTCCGCTCAAGCTGGGCTGTGTGACGAA	428
OY	1579	cccccggttcaacccgcgcgcgtgcgccttatccggttaactatggtcgttgaatcaacc	1637
Db	429	CCCCGTTTCACCCGACCCGCTGCCTCTTATCCGGGTATCTATCGTTGAGTCCAAACCCGG	488
OY	1638	gtaagacacgaactatcgccactgacgacgacgacactgtaacagatlaacagacgag	1697
Db	489	GTAAGACACGACTTATCGCCACTGCGACGACCCACTGCTAACAGATTTACAGAGCGAGG	548
OY	1698	tatgtaggcgtgtcctacagagttccttgaagtgtgtgcctaaactaaggctcaactaaga	1757
Db	549	TATGTAGCGCGGTGCTACANAGTTCTTGAAAGTGTGGCGCTTAACCTACGCTCACAGAGA	608
OY	1758	acagta-cttggtatctbgtgcgtcgtgcgtgaagcaggttaactcttgcggaaaaa--	1814
Db	609	ACAGTATTTGTTGATCTGCGCTGTGTAACCAAGTTACCTTCGGAAAAAAGANGTGGGT	668
OY	1815	agctcttg-----atccggaacaacaacacccgcgtgtagcgtgtgtttcttggtca	1870
Db	669	AGCCTTGTGATTCGGCNAAAAAAACCAACCGCGTGTAGAGTGTGTTTTTTNTTNC	728
OY	1871	gcagca 1876	
Db	729	AAAGCA 734	
RESULT	11		
AG010489			
LOCUS	718 bp	DNA	GSS 14-APR-1999
DEFINITION	Homo sapiens genomic DNA, 21q region, clone: P7H8X25, genomic		
ACCESSION	AG010489 AG004363		
VERSION	AG010489.1 GI:3327308		
KEYWORDS	GSS.		
SOURCE	Homo sapiens DNA, clone:P7H8X25.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryotes: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:		
AUTHORS	Mammalia: Eutheria: Primates: Catarrhini: Homiidae: Homo.		
TITLE	1 (bases 1 to 718)		
JOURNAL	Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.		
REFERENCE	Homo sapiens genomic DNA, chromosome 21q		
AUTHORS	Published only in DataBase (1998) In press		
JOURNAL	2 (bases 1 to 718)		
TITLE	Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.		
REFERENCE	Direct Submission		
AUTHORS	Submitted (16-JUL-1998) to the DDBJ/EMBL/Genbank databases.		
JOURNAL	Maehara Hattori, Kitasato University, Department of Science, Jpn		
REFERENCE	Sequencing Laboratory: Kitasato 1-15-1, Sagamihara 228, Japan		
AUTHORS	(E-mail:hattori@hgc.ims.u-tokyo.ac.jp, Tel:0427-78-9732,		
JOURNAL	Fax:0427-78-9561)		

COMMENT	On Feb 5, 1999 this sequence version replaced gi:2789511.
AG004363: Submitted (17-Jan-1998).	
FEATURES	
source	Location/Qualifiers
	1..718
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/chromosome="21"
	/clone="F7H8x25"
	/map="21q"
BASE COUNT	176 a 191 c 187 g 158 t 6 others
ORIGIN	
Query Match	7.6%; Score 615.2; DB 219; Length 718;
Best Local Similarity	97.2%; Pred. No. 1.3e-160;
Matches 667; Conservative	0; Mismatches 15; Indels 4; Gaps 4.
Db	1224 tccacagatcgggagatgaacgacgaagaacatgtgacaaagccagcaaaagcc 1283
	21 tccccggagagagggatattaccgacgaagaaacatgtgacaaagccagcaaaagcc 80
Qy	1284 aggaacccgtaaaaagagccgctgtgtcgtgttccataggtctccgccccctgacgag 1343
Db	81 AGGAACCCGTA AAAAGGCCGCTTGCGTGGCGGTTTTCATATAGGCTCCGCCCTTGACGAG 140
Qy	1344 catcaaaaaatcgcagcgtccaagtcaagagtgtgacgaacccgacaggtatataagatgc 1403
Db	141 CATCACAAAAATTCGACGCTCAAGTCAGAGGTGGCGAAGCCGACAGAGATATPAAGATATC 200
Qy	1404 caggcgcttcccccttggaaagctccctcgtgagctccctgttccgacctgcgcgtttacc 1463
Db	201 CAGGCGTTTCCCCCTGGAAAGCTCCCTCGTGGCTCTCCCTTCACACCTGGCGCTATTAC 260
Qy	1464 ggaatacctgtccgccttctcccttcgaggaagcgttgcgcttccatgtctacgcgtgt 1523
Db	261 GGATACCTGTCCGCGCTTCTCCCTTCGGGAACGTTGCGCTTTCATATAGCTACACCTGT 320
Qy	1524 aagtatctaatgttggttgaagtcggttcgtctccaagcttgggtgtgtgacgaacccccc 1583
Db	321 AGGTATCTCAGTTCGGTGTGAGTGTGCTTCGCTCCAAAGCTGGGCTGTGTGCACGAACCCCCC 380
Qy	1584 gtccaccccgacgcgtcgtgcgctctaccgcgttaactatcgtcttgagtcacacccggtaa 1643
Db	381 GTTTCAGCCCGACCGCTGGCGCTTATCCGGTAACTATCGTCTTGAGTGCACACCCGGTAA 440
Qy	1644 caagactatcgcacactggaacgacgaccactgttlaacagattagcaagcgaagtatgta 1703
Db	441 CAGGACTTATGCCACTGGCAGCAGCACTGGTAAACAGATTAGCAGACNNAGTATGTA 500
Qy	1704 ggcggtgtc-acaaagttcttgaagtgtgtgccttaactagcgtc-acactagaagaagc 1761
Db	501 GCGCGGTCTAACAGAGTTCTTGAAGTGGGGCTTAATCTACGGCTTAACATTAAGAAGACG 560
Qy	1762 tatlttgatctgcgcctcgtcgtgaagccagttacc-ttcggaaaaaagtttgtatcct 1820
Db	561 TATTTTGATCTGCGCTCTGCTGTAAGCCAGTTANCTTTGGAAAAAGATTGGTACTCT 620
Qy	1821 tgatccggaacaaacacacccgctgtgtagcgttggttttttgttgcgaagcagcagatt 1880
Db	621 TGATCGGGCAAAACAACACCGCTGGTATC-GTGTTTTTTTGTTTTGCAGAGCAGATATA 679
Qy	1881 acgcgcagaaaaaaaggatctcaaga 1906
Db	680 ACGCGCAGAAAAAAGGATCTCAAGA 705
RESULT 12	
AU001481	687 bp mRNA EST 15-JAN-1999
LOCUS	AU001481 Bombyx mori p50(Daizo) Bombyx mori cDNA clone Fbm0478f.
DEFINITION	mRNA sequence.
ACCESSION	AU001481
VERSION	AU001481.1 GI:4157725

KEYWORDS EST.
 SOURCE domestic silkworm.
 ORGANISM Bombyx mori
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia
 ; Bombycoidea; Bombycidae; Bombyx.
 REFERENCE 1 (bases 1 to 687)
 AUTHORS Mita, K., Morino, M., Shimada, T., Okano, K. and Maeda, S.
 TITLE Establishment of cDNA database of Bombyx mori
 JOURNAL Unpublished (1999)
 COMMENT Contact: Mita K
 Genome Research Group
 National Institute of Radiological Sciences
 Anagawa 4-9-1, Inage, Chiba 263-8555, Japan
 Email: kmita@nirs.go.jp
 PROJECT - GREST project by JST.
 Location/Qualifiers
 1. 687
 /organism="Bombyx mori"
 /strain="p50(Daizo)"
 /db_xref="taxon:7091"
 /clone="Tbm0478f"
 /clone_lib="Bombyx mori p50(Daizo)"
 BASE COUNT 145 a 209 c 183 g 149 t 1 others
 ORIGIN
 Query Match 7.5%; Score 606.8; DB 106; Length 687;
 Best Local Similarity 99.5%; Pred. No. 2.9e-158;
 Matches 608; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 Oy 1076 ccagcgcatatgaatgcgcaacgcgagcgagcggttgcttgatggcgctc 1135
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 77 CCAGGTCATTAATGAATCGCCACGCGGAGAGGCGTTGGTATGGCGCTC 136
 Oy 1136 ttccgcttcctcgctactgaactgcgtgcgtgcgttcgctgcgagcgatc 1195
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 137 TTCGGCTTCCTCGCTCACTGACTCGCTGCGCTGCTGCTGCGGAGCGGATC 196
 Oy 1196 agctcaactcaagcggttaatacagttatcacagaatcacaggagataacgcaagaa 1255
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 197 AGCTCACTCAAAAGCGGTAATACGGTTATCCACAGAAATCAGGGGATTAACGAGAAAGAA 256
 Oy 1256 catgtagcaaaaagccagcaaaaagccagaaacccgttaaaaagccggttcgttcgagtc 1315
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 257 CATGTAGCAAAAAGCGCAGCAAAAAGCGCAGAAACCGTAAAGGCCCGCTTGGCGCTN 316
 Oy 1316 ttccatagagctcgccccctcagcagcatcaaaaatcgaacgtcaagtcaagatg 1375
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 317 TTTCATAGGCTCGCCCTCCTACGAGCATCCAAAAATCGACGCTCAAGTCAGAGTG 376
 Oy 1376 gcgaaacccgacagactataaagataccagcgcttcccccctggaagctccctcgtcg 1435
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 377 GCGAAACCGACAGGACTATAAAGATACGAGGCTTCCCTCGGAAGCTCCCTCGTGGC 436
 Oy 1436 ctctcctgttcgacccctcgcttacccgataccctgtccgcttcccttcgggaag 1495
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 437 CTCTCTCTTCCACACCTCGCCCTTACCGGATACCTGTCCGCTTCTCCCTTCGGGAG 496
 Oy 1496 cgtggagcttcccaatgcacgctgtagtattccaafttggttgaattgttcgctc 1555
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 497 CGTGGGCTTTTCATAGCTCACGCTGTAGGTATCTCAGTTGAGTGTGCTTCCTC 556
 Oy 1556 caagctgggctgtgtgcaagaaaccccgcttcagcccgacgctgcgcttaccgta 1615
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 557 CAAGCTGGGGTGTGTGACAGAACCCCGGTTCAGCCGACCGTGGCGCTTATCCGGTAA 616
 Oy 1616 ctatcgtcttgatccaaaccggttaagaacgactatcgccactggcagcagccactg 1675
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 617 CTATCTCTTTGAATCAACCGGTTAGACACGACTTATCCCTACGACAGCACTGG 676
 Oy 1676 taacagatt 1686
 |||||||||||

Db 677 TAACAGGATTA 687
 RESULT 13
 LOCUS AU081044 1163 bp mRNA EST 15-NOV-1999
 DEFINITION AU081044 Oncorhynchus mykiss Kidney infected by infectious
 hematopoietic necrosis virus Oncorhynchus mykiss cDNA clone KBL,
 mRNA sequence.
 ACCESSION AU081044
 VERSION AU081044.1 GI:6431392
 KEYWORDS EST.
 SOURCE rainbow trout.
 ORGANISM Oncorhynchus mykiss
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
 REFERENCE 1 (bases 1 to 1163)
 AUTHORS Sakai, M. and Kono, T.
 TITLE The EST analysis of kidney and gill cells rainbow trout,
 Oncorhynchus mykiss, infected with infectious hematopoietic
 necrosis virus
 JOURNAL Unpublished (1999)
 COMMENT Contact: Masahiro Sakai
 Faculty of Agriculture
 Miyazaki University
 1-1 nishi gakuenkibanaadai, Miyazaki, Miyazaki 889-2192, Japan
 Email: aob208@cc.miyazaki-u.ac.jp.
 Location/Qualifiers
 1. 1163
 /organism="Oncorhynchus mykiss"
 /db_xref="taxon:8022"
 /clone="KBL"
 /clone_lib="Oncorhynchus mykiss Kidney infected by
 infectious hematopoietic necrosis virus"
 /tissue_type="Kidney infected by infectious hematopoietic
 necrosis virus"
 BASE COUNT 272 a 307 c 298 g 286 t
 ORIGIN
 Query Match 7.4%; Score 604.2; DB 107; Length 1163;
 Best Local Similarity 99.4%; Pred. No. 1.9e-157;
 Matches 617; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
 Oy 1076 ccagctgcatatgaatgcgcaacgcgagcgagcggttgcttgatggcgctc 1135
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 544 CCAGGTCATTAATGAATCGCCACGCGGAGAGGCGTTGGTATGGCGCTC 603
 Oy 1136 ttccgcttcctcgctactgaactgcgtgcgttcgctgcgtgcgagcgatc 1195
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 604 TTCGGCTTCCTCGCTCACTGACTCGCTGCGCTGCTGCTGCGGAGCGGATC 663
 Oy 1196 agctcaactcaagcggttaatacagttatcacagaatcacaggagataacgcaagaa 1255
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 664 AGCTCACTCAAAAGCGGTAATACGGTTATCCACAGAAATCAGGGGATTAACGAGAAAGAA 723
 Oy 1256 catgtagcaaaaagccagcaaaaagccagaaacccgttaaaaagccggttcgttcgagtc 1315
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 724 CATGTAGCAAAAAGCGCAGCAAAAAGCGCAGAAACCGTAAAGGCCCGCTTGGCGGCTT 783
 Oy 1316 ttccatagagctcgccccctcagcagcatcaaaaatcgaacgtcaagtcaagatg 1375
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 784 TTTCATAGGCTCGCCCTCCTACGAGCATCCAAAAATCGACGCTCAAGTCAGAGTG 843
 Oy 1376 gcgaaacccgacagactataaagataccagcgcttcccccctggaagctccctcgtcg 1435
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 844 GCGAAACCGACAGGACTATAAAGATACAGGCTTTCCTCGGAAGCTCCCTCGTGGC 903
 Oy 1436 ctctcctgttcgacccctcgcttacccgataccctgtccgcttcccttcgggaag 1495
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 904 CTCTCTGTTCCGACCCCTGCGCTTACCGGATACCTGTCCGCTTCTCCCTTCGGGAG 963

QY 1496 cgtgagcgttcacaaatcagcgtgtaggtatcagttcggtgtagtgcgtc 1555
 Db 964 CGTGGCGCTTTCATACCTACGCTGAGTATCTAGTTCGTTGAGTTCGCTC 1023
 QY 1556 caactggtggtgtgtgacgaacccccgttcagccgacgctgagcctatccggtaa 1615
 Db 1024 CAACCTGGCGGTGTGTCAGAA-CCCCGTTACGCCACCGCTGCTTATCCGGTAA 1082
 QY 1616 ctatcgcttgaatccaaacccggttaagacacagactatcgcacactgagcagacactgg 1675
 Db 1083 CTATCGCTTTCATACCTACGCTGAGTATCTAGTTCGTTGAGTTCGCTC 1142
 QY 1676 taacagattagcagcagcag 1696
 Db 1143 TAACAGATTAGCAGAGCGAG 1163

RESULT 14
 AJ281699 636 bp mRNA EST 30-JUN-2000
 LOCUS 4A3A-P9E3-F Anopheles gambiae immune competent 4A3A Anopheles
 DEFINITION gambiae cDNA clone 4A3A-P9E3, mRNA sequence.
 ACCESSION AJ281699
 VERSION AJ281699.1 GI:6929578
 KEYWORDS EST.
 SOURCE African malaria mosquito.
 ORGANISM Anopheles gambiae
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae
 ; Anopheles.
 1 (bases 1 to 636)
 Dimopoulos,G., Casavant,T.L., Chang,S., Scheetz,T., Roberts,C.,
 Donohue,M., Schultz,J., Benes,V., Bork,P., Ansong,W., Soares,M.B.
 and Kafatos,F.C.
 Anopheles gambiae pilot gene discovery project: identification of
 mosquito innate immunity genes from expressed sequence tags
 generated from immune-competent cell lines
 Proc. Natl. Acad. Sci. U.S.A. 97 (12), 6619-6624 (2000)
 20300950
 COMMENT Contact: Dimopoulos G
 Fotis C. Kafatos Laboratory
 European Molecular Biology Laboratory
 Meyerhofstrasse 1, 69117 Heidelberg, Germany.
 LOCATION/Qualifiers
 1. .636
 /organism="Anopheles gambiae"
 /strain="4A r/r"
 /db_xref="taxon:7165"
 /clone="4A3A-P9E3"
 /clone_1id="Anopheles gambiae immune competent 4A3A"
 /cell_line="immune competent 4A3A"
 /lab_host="E. coli DH10B"
 /note="Vector: pT773D-Pac (Pharmacia) with a modified
 polylinker; Site_1: EcoRI; Site_2: NotI; sequenced from
 forward priming site which reads from the 3' end of the
 cDNA. The 4A3A is a directionally cloned and normalized
 cDNA library that was constructed from the 4A3A cell line
 oligo-T primed cDNA according to: Bernaldo, Lennon & Soares
 (1986) : Normalization and Subtraction: Two approaches to
 Facilitate Gene Discovery, Genome Research 6, 791-806."

BASE COUNT 152 a 179 c 166 g 139 t
 ORIGIN

Query Match 7.4%; Score 604; DB 104; Length 636;
 Best Local Similarity 98.9%; Pred. No. 1,7e-157;
 Matches 629; Conservative 0; Mismatches 5; Indels 2; Gaps 2;

QY 1202 ctcaaaagcggtataatgattacagaaatcagggatagcaggaagaagaacatg 1261
 Db 1 CTCAAAAGCGGTATATCGTTATCCACAGATCAGGGGATTAACGGAAGAAACATGTG 60

QY 1262 agcaaaagccagaaagcaggaacgtaaaagcgcggttgcgtgtttcca 1321

Db 61 AGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAGGCCGGGTGTCGCTTTTCCA 120
 QY 1322 tagctccgcccccttgacgagcatcacaaaaatcagaatcagaatgaggtggcga 1381
 Db 121 TAGGCTCGCCCCCTTGACGAGCATACAAAATGACAGCTAAATGACAGAGTGCGA 180
 QY 1382 cccagcagactataataacacagcgcttcccccctggaagctccctgtgcgtctc 1441
 Db 181 CCGGACAGAGACTATAAATATACAGAGCGTTTCCCTGGAAGCTCCCTGTCGCTCC 240
 QY 1442 tgttcgaacctgcccgtctaacggaatccctgtgcgcttcccttcgggaagcgtg 1501
 Db 241 TGTTCGACCCCTGCGCTTACCGGATACCTGTCGCTTTCCTTCGCGGAAGCTGCG 300
 QY 1502 gcttctcaatgctacgctgtaggtatcgaatgctggtgtaggtcgttgcctcaagct 1561
 Db 301 GCTTCTCATAGCTACCGCTGAGGTATCTCAGTTGCGGTGAGTGTGCTCCAGCT 360
 QY 1562 gggctgtgtgcaagaccgccgttcagccgacccgctgcgctatccgtaactatcg 1621
 Db 361 GGGCTGTGTGACAGAACCCCGCTTACAGCCGACCGCGCTTATCGGTAACCTATCG 420
 QY 1622 tcttgatccaaacccggtgaagacacgactatcgccactgagcagcactggttaacg 1681
 Db 421 TCTTGATCCAAACCCGTTAGACACGACTTATCCGCTGCGAGAGCCACTGTGTAACG 480
 QY 1682 gattgagcagcaggaattgtgtagcggtgtacagaatcttgaatggtgtggtc-taact 1740
 Db 481 GATTAGCAGAGGAGGATGATGAGCGGTGTACAGAGTTCTTGAAGTGTGCGCTTTTCT 540
 QY 1741 acggtacactagaagcagcagattgttatctgcgtctctggaagcagttactctg 1800
 Db 541 ACGGCTACACTAGAGAACACAGATTGTGATCGCTGCTGTAAGCAGTTACTCTCG 600

QY 1801 g-aaaaagattgtagctcttgatccggcaacaa 1835
 Db 601 GAAAAAAGAGTTGATGCTTGTGATCGCGCAACAA 636

RESULT 15
 A0876119 784 bp DNA GSS 08-NOV-1999
 LOCUS A0876119/c
 DEFINITION V133E1 mtn-3xHA/LacZ insertion library, strain Y2278 Saccharomyces
 cerevisiae genomic 5', DNA sequence.
 ACCESSION A0876119
 VERSION A0876119.1 GI:6288363
 KEYWORDS GSS.
 SOURCE Saccharomyces cerevisiae.
 ORGANISM Saccharomyces cerevisiae
 Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 Saccharomycetaceae; Saccharomycetaceae; Saccharomycetes.
 1 (bases 1 to 784)
 Ross-Macdonald,P., Roemer,T., Coelho,P.S.R., Agarwal,S., Kumar,A.,
 deSilva,S.A., Cheung,K.H., Sheehan,A., Symoniatidis,D., Jansen,R.,
 Umekang,L., Heldman,M., Nelson,K., Iwasaki,H., Kanada,D., Lugo,R.,
 Hager,K., Miller,P., Roeder,G.S. and Snyder,M.
 Large-Scale Analysis of the Yeast Genome by Transposon Tagging and
 Gene Disruption
 Unpublished (1999)
 JOURNAL Contact: Kumar A
 COMMENT Michael Snyder, Dept. of Mol. Cell. and Dev. Biology
 Yale University
 P.O. Box 208103, New Haven, CT 06520-8103, USA
 Tel: 203 432 9949
 Fax: 203 432 6161
 Email: anuj.kumar@yale.edu
 te of mtn-3xHA/LacZ insertion.
 Seq primer: GGCCTCTTCTTGTGGAAGTAC
 Class: transposon-tagged.
 Location/Qualifiers
 1. .784
 /organism="Saccharomyces cerevisiae"


```

/strain="Y2278 - S288C background, cir(+)"
/db.xref="taxon.4933"
/clone.lib="mtn-3XHA/lacZ Insertion Library, strain Y2278
/lab.host="E. coli"
/note="Vector: pHS6-Sal; A yeast genomic DNA library
without 2 micron or mitochondrial DNA was prepared in
pHS6-Sal; genomic DNA was size-fractionated (DNA of
roughly 2-3 kb in length) prior to cloning. This library
was subsequently mutagenised with a mtn-3XHA/lacZ
multitransposon containing lacZ, URA3, and tet resistance.

```

BASE COUNT	187 a	200 c	210 g	186 t	1 others
ORIGIN					

Query Match	7.48;	Score 602.6;	DB 235;	Length 784;
Best Local Similarity	97.68;	Pred. NO. 4.5e-157;		
Matches 611;	Conservative 0;	Mismatches 15;	Indels 0;	Gaps 0;

OY	1304	gtctctgagctttttccatagagctccgcccccttgacagagcaacaaatccagcgtc	136
Db	784	ggtctgtggcgttttttccattagcctcgcccccttgacagagcaacaaatccagcgtc	725
OY	1364	aagtcagaagcttgagcaaaaccggaacagactataaagataccagagcgcttccccctgagag	1423
Db	724	aagtcagaagcttgagcaaaaccggaacagactatpaaagatccagagcggttccccctgagag	665
OY	1424	ctccctcgttgagctctccctgtttccgaacctgcgcgttacccgatatctgtccgcttctc	1483
Db	664	ctccctcgttgagctctccctgtttccgaacctgcgcgttacccgatatctgtccgcttctc	605
OY	1484	ccctctgggaagcgttgagcgctttctcaabgctcaacgctgtaagatctcaagttcggtgta	1543
Db	604	ccctctgggaagcgttgagcgctttctcaabgctcaacgctttagatctcaagttcggtgta	545
OY	1544	ggtcgtctgcctcaagcttgagcgtctgtgtgaagaaccccccgcttacagccgacgctgcgc	1603
Db	544	ggtcgtctgcctcaagcttgagcgtctgtgtgaagaacccccctcgttacagccgacgctgcgc	485
OY	1604	ctcttcgcgtaacatctgcttttgaagccaaaccggaagaacacgacttatgcacctgac	1663
Db	484	ctcttcgcgtaacatctgcttttgaagccaaaccggaagaacacgacttatgccacctgcac	425
OY	1664	agcagccacctggtlaacagagcttagacagagcagagtatgtgaagcgtgtgtacagaattct	1723
Db	424	agcagccacctggtlaacagagatttagacagagcagagatgtatgaagcgggtgcctacagattctt	365
OY	1724	gaagctggttgagcctaactacagctacacatagaagacagatatttgatctgcgtctgcgt	1783
Db	364	gaagctggttgagcctaactacagctacacatagaagacagatatttgatctgcgtctgcgt	305
OY	1784	gaagccaggttaaccttggaanaaagatttgtagctcttcatctcggcaacaacaacacgcg	1843
Db	304	gaagccaggttaaccttggaanaaagatttgtagctcttcatctcggcaacaacaacacacgcg	245
OY	1844	tggttagcaggttgattttttgtttgttgcagaagcagcagatctacgcgcaagaaaaaagatctca	1903
Db	244	tggttagcaggttgattttttgtttgttgcagaagcagcagatttagccgcaaaaaaagatctca	185
OY	1904	agaagaatcccttgatctcttctacagc	1929
Db	184	agaagaatcccttgatctcttctctactg	159

Search completed: November 3, 2001, 17:35:03
Job time: 7989 sec

THIS PAGE BLANK (USPTO)